Universität Tübingen

Prof. Victor Batyrev

LINEARE ALGEBRA I

Wintersemester 20015/2016

Mathematisches Institut

Inhaltsverzeichnis

1 Grundlegende Begriffe

Bezeichnungen:
Summe:

$$\sum_{i=1}^{k} a_i := a_1 + a_2 + \cdots + a_k$$

Produkt:

$$\prod_{i=1}^{k} a_i := a_1 \cdot a_2 \cdots a_k.$$

Beispiel:

$$n! = \prod_{k=1}^{n} k.$$

Operationen mit Mengen

$$A \cap B := \{x \mid x \in A \text{ und } x \in B\},$$

Differenz von Mengen $A \setminus B$
Die Gleichungen

$$A \cap (B \cup C) = (A \cup B) \cap (A \cup C).$$

$$A \cup (B \cap C) = (A \cap B) \cup (A \cap C).$$

Durschnitt und Vereinigung von Mengen und Mengenfamilien

$$\bigcap_{i \in I} M_i := \{x \mid x \in M_i \ \ \forall i \in I\},$$

$$\bigcup_{i \in I} M_i := \{x \mid \exists i \in I \text{ mit } x \in M_i\}.$$

2 Vollständige Induktion

3 Lineare Gleichungssysteme

3.1 Matrix von LGS

Ein **lineares Gleichungssystem** (kurz **LGS**) mit m Gleichungen und n **Variablen** $x_1, \ldots, x_n$ (man nennt die auch **Unbestimmten** oder **Unbekannten**) hat folgendes Aussehen:

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{gathered}$$

Die Koeffizienten a_{ij} des LGS schreibt man als rechteckiges Schema

$$A := (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & a_{m2} & a_{mn} \end{pmatrix}$$

welches die **Koeffizientenmatrix von LGS** genannt wird.
Die Koeffizienten b_i des LGS schreibt man als Spalte

$$b := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Das rechteckige Schema

$$(A, b) := \begin{pmatrix} a_{11} & a_{12} & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{2n} & b_2 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & a_{m2} & a_{mn} & b_m \end{pmatrix}$$

heißt die **erweiterte Koeffizientenmatrix von LGS**.
LGS heißt **homogen** $\Leftrightarrow b = 0$;
LGS heißt **inhomogen** $\Leftrightarrow b \neq 0$;
LGS heißt **quadratisch** $\Leftrightarrow m = n$;
Beispiel von LGS:

$$x_1 + 2x_2 + 3x_3 = 6$$

$$x_1 - 2x_2 + x_3 = 0$$

$$2x_2 - 3X_3 = -1.$$

Die erweiterte Matrix von LGS:

$$(A, b) = \begin{pmatrix} 1 & 2 & 3 & 6 \\ 1 & -2 & 1 & \\ 0 & 2 & -3 & -1 \end{pmatrix}$$

Definition 3.1 Mit $\text{Lös}(A, b) \subset \mathbb{R}^n$ bezeichnen wir die Lösungsmenge des LGS mit der erweiterte Matrix (A, b).

Für das Beispiel gilt: $\text{Lös}(A, b) = \{(1, 1, 1)\}$.

3.2 Der Raum $\mathbb{R}^n$

Definition 3.2 Der **reelle n-dimensionale Standartraum** $\mathbb{R}^n$ ist die Menge der geordneten n-Tupel von reellen Zahlen, d.h.

$$\mathbb{R}^n := \{x = (x_1, \ldots, x_n) \ : \ x_1, \ldots, x_n \in \mathbb{R}\}.$$

Die Elemente $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ nennt man **Vektoren**. Die Zahlen $x_1, \ldots, x_n$ heißen **Komponenten** von x. Ein Vektor $(0, \ldots, 0) \in \mathbb{R}^n$, dessen alle Komponenten gleich Null sind, wird **Nullvektor** genannt und einfach mit 0 bezeichnet.

Operationen mit Vektoren aus $\mathbb{R}^n$:
1) **Addition**:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n);$$

2) **Multiplikation mit einer Zahl** $\lambda \in \mathbb{R}$:

$$\lambda \cdot (x_1, \ldots, x_n) := (\lambda \cdot x_1, \ldots, \lambda \cdot x_n).$$

Definition 3.3 Das **Skalarprodukt** $\langle x, y \rangle$ zweier Vektoren $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ ist durch die folgende Formel definiert:

$$\langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

Summenzeichen
(Vektoren x, y in $\mathbb{R}^2$ oder in $\mathbb{R}^3$ bezeichnet man oft mit $\vec{x}, \vec{y}$ und das Skalarprodukt $\langle x, y \rangle$ mit $\vec{x} \cdot \vec{y}$.)

Der Beweis der folgenden Eigenschaften des Skalarprodukts wird dem Leser überlaßen:

Proposition 3.4 *Für alle $x, y, x', y' \in \mathbb{R}$ und $\lambda \in \mathbb{R}$ gelten die folgenden Gleichungen:*
(i) $\langle x, y\rangle = \langle y, x\rangle$;
(ii) $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$;
(iii) $\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle$;
(iv) $\langle \lambda \cdot x, y\rangle = \langle x, \lambda \cdot y\rangle = \lambda \cdot \langle x, y\rangle$.

Eine Matrix $A = (a_{ij})$ mit n Spalten und m Zeilen nennen wir $m \times n$-**Matrix**. Mit

$$Z_i(A) := (a_{i1}, a_{i2}, \ldots, a_{in})$$

bezeichnen wir die i-te Zeile von A, $i \in \{1, 2, \ldots, m\}$, und mit

$$S_j(A) := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

die j-te Spalte von A. Mit diesen Bezeichnungen kann man LGS kurz als Gleichungssystem

$$\langle Z_1(A), x\rangle = b_1, \quad \langle Z_2(A), x\rangle = b_2, \quad \ldots, \quad \langle Z_m(A), x\rangle = b_m,$$

oder noch kurzer als

$$\langle Z_i(A), x\rangle = b_i, \;\; i \in \{1, \ldots, m\},$$

aufschreiben, wobei $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ein unbekannter Vektor ist, der gegebene Skalarprodukte $b_1, \ldots, b_m$ mit Zeilen-Vektoren $Z_1(A), \ldots, Z_n(A) \in \mathbb{R}^n$ hat.
Andererseits kann man das gleiche System als eine Vektor-Gleichung

$$x_1 S_1(A) + x_2 S_2(A) + \cdots + x_n S_n(A) = b$$

für Koeffizienten $x_1, \ldots, x_n$ bei Spalten-Vektoren $S_1(A), S_2(A), \ldots, S_n(A), b \in \mathbb{R}^m$ interpretieren.

3.3 Elementare Zeilenoperationen

Definition 3.5 Die folgenden drei Operationen mit Zeilen einer Matrix A werden **elementare Zeilenumformungen** genannt:
Typ 1: Vertauschung der i-ten und der j-ten Zeilen ($i \neq j$);
Typ 2: Addition der λ-fachen j-ten Zeile zur i-ten Zeile ($i \neq j, \lambda \in \mathbb{R}$);
Type 3: Multiplikation der i-ten Zeile mit dem Skalar $\lambda \in \mathbb{R}$ $(\lambda \neq 0)$.

Ist A' die durch eine elementare Zeilenumformung entstandene Matrix, so gilt
Typ 1: $Z_k(A') = Z_k(A)\ \forall k \neq i, j$ und $Z_i(A') = Z_j(A)$, $Z_j(A') = Z_i(A)$;
Typ 2: $Z_k(A') = Z_k(A)\ \forall k \neq i$ und $Z_i(A') = Z_i(A) + \lambda Z_j(A)$;
Typ 3: $Z_k(A') = Z_k(A)\ \forall k \neq i$ und $Z_i(A') = \lambda Z_i(A)$;

Satz 3.6 *Sei (A, b) die erweiterte Koeffizientenmatrix eines LGS. Ist (A', b') aus (A, b) durch endlich viele elementare Zeilnumformungen entstandene Matrix, so gilt*

$$\text{Lös}(A, b) = \text{Lös}(A', b')$$

.

Beweis. Es genügt zu beweisen, dass die Lösungsmenge bei einer einzigen elementaren Zeilenumformung unverändert bleibt, denn dann ändert die Wiederholung nichts. Bei einer elementaren Zeilenumformung vom Typ 1 diese Eigenschaft is offensichtlich, denn die Reihenfolge der Gleichungen ist für die Lösunsmenge gleichgültig. Bei einer elementaren Zeilenumformung vom Typ 2 oder vom Typ 3 zeigen wir zunächst, dass

$$\text{Lös}(A, b) \subseteq \text{Lös}(A', b').$$

Ist $c = (c_1, \ldots, c_n) \in \text{Lös}(A, b)$ (d. h. $\langle Z_k(A), c\rangle = b_k$, $\forall k \in \{1, 2, \ldots, m\}$), so ist $\langle Z_k(A'), c\rangle = \langle Z_k(A), c\rangle = b_k = b'_k$, falls $k \neq i$. Bei einer elementaren Zeilenumformung vom Typ 2 hat man:

$$\langle Z_i(A'), c\rangle = \langle Z_i(A) + \lambda Z_j(A), c\rangle = \langle Z_i(A), c\rangle + \langle \lambda Z_j(A), c\rangle =$$

$$= \langle Z_i(A), c\rangle + \lambda\langle Z_j(A), c\rangle = b_i + \lambda b_j = b'_i$$

Bei einer elementaren Zeilenumformung vom Typ 3 hat man:

$$\langle Z_i(A'), c\rangle = \langle \lambda Z_i(A), c\rangle = \lambda\langle Z_i(A), c\rangle = \lambda b_i = b'_i.$$

Die umgekehrte Inklusion

$$\text{Lös}(A', b') \subseteq \text{Lös}(A, b)$$

folgt aus der Tatsache, dass die Matrix (A, b) selbe aus (A', b') durch eine elementare vom Typ 2 (oder vom Typ 3) entsteht.

□

3.4 Das Eliminationsverfahren von Gauss

Definition 3.7 Eine $m \times n$ Matrix $A = (a_{ij})$ heißt in **Zeilenstufenform**, wenn folgendes gilt:
(i) es gibt eine Zahl r mit $0 \leq r \leq m$, so dass $Z_{r+1}(A) = \cdots = Z_m(A) = 0$;
(ii) ist $r \geq 1$, so gibt es Zahlen $j_1, \ldots, j_r$ mit $1 \leq j_1 < \cdots < j_r \leq n$, so dass

$$a_{1j_1} \neq 0, \ldots, a_{rj_r} \neq 0$$

und

$$a_{kl} = 0 \quad \text{für alle } k, l \text{ mit } k \leq r \text{ und } l < j_k.$$

Eine Matrix A in Zeilenstufenform heißt in **normierter Zeilenstufenform**, wenn zusätzlich gilt

$$a_{1j_1} = \cdots = a_{rj_r} = 1$$

und

$$a_{k,j_l} = 0 \quad \text{für all } k, l \text{ mit } k < l \text{ und } 1 \leq l \leq r.$$

Die Zahl r (das heißt die **Anzahl von Null verschiedenen Zeilen**) einer Matrix A in Zeilenstufenform wird auch mit $r(A)$ bezeichnet.

Satz 3.8 *Jede $m \times n$-Matrix A kann man durch elementare Zeilenumformungen vom Typ 1 und 2 in eine Matrix B in Zeilenstufenform überführen.*

Beweis. Wir beweisen diese Aussage mit Induktion über die Summe $m + n$. Ist $m = n = 1$ und damit $m + n = 2$, so ist A schon in Zeilenstufenform und wir brauchen nichts zu beweisen.
Nach Induktionsvoraussetzung können wir annehmen, dass die Aussage des Satzes für alle n, m mit $n + m \leq k$ gilt. Sei A eine $m \times n$-Matrix mit $k + 1 = m + n$.
Ist die erste Spalte $S_1(A)$ ein Null-Vektor, so ist kann man nach der Induktionsvoraussetzung die Matrix A', die aus A durch Streichen der ersten Spalte entsteht, mit Elementarumvormungen in Zeilenstufenform überführen. Die gleichen Elemenentarumvormungen mit Matrix A werden auch die Matrix A in eine Zeilenstufenform überführen.
Ist die erste Spalte $S_1(A)$ kein Null-Vektor, so können wir durch Vertauschung von Zeilen eine Matrix A' erreichen, so dass $a'_{11} \neq 0$. Durch Additionen der $(-a_{i1}/a'_{11})$-fachen ersten Zeile von A' zur i-ten Zeile von A' erreichen wir eine Matrix A'', so dass alle Einträge a''_{i1} der ersten Spalte $S_1(A'')$ ausser a''_{11} gleich Null sind. Durch Streichen der ersten Zeile der Matrix A'' bekommen wir eine Matrix A''', so dass die erste Spalte $S_1(A''')$ ein Null-Vektor ist. Nach der Induktionsvoraussetzung kann

man die Matrix A''' mit Elementarumformungen von Zeilen in die Zeilenstufenform bringen. Damit überführen wir auch die Matrix A in die Zeilenstufenform. □

Bemerkung 3.9 Man merkt es leicht, daß für eine erweiterte Koeffizienten Matrix (A, b) des LGS in Zeilenstufenform gilt:
Ist $1 \leq j \leq r$, so besitzt die lineare Gleichung $\langle Z_j(A), x\rangle = b_j$ keine Variabel x_i mit $i < j$. Deswegen nennt man das Verfahren, welches eine Koeffizientenmatrix (A, b) in eine Zeilenstufenform überfuhrt, das **Eliminationsverfahren** (dadurch eliminiert man Variablen in LGS), oder **Eliminationsverfahren von Gauss**.

Satz 3.10 *Jede Matrix A in Zeilenstufenform kann man durch elementare Zeilenumformungen vom Typ 2 und 3 in eine Matrix B in normierter Zeilenstufenform überführen.*

Beweis. Zuerst multiplizieren wir jede k-te Zeile von A mit $1/a_{kj_k}$, $1 \leq k \leq r$. Damit erhalten wir eine Matrix A' in Zeilenstufenform mit

$$a'_{1j_1} = \cdots = a'_{rj_r} = 1.$$

Nun beweisen wir die Aussage des Satzes mit Induktion über die Zahl r, die der Anzahl von Null verschiedenen Zeilen in der Matrix A' gleich ist.
Ist $r = 1$, so gibt es nichts zu beweisen. Nehmen wir an, dass die Aussage des Satzes für ein $r = k \geq 1$ gilt.
Sei $r = k + 1$. Wir addieren die $(-a'_{kj_r})$-fache r-te Zeile von A' zur k-ten Zeile von A', $1 \leq k \leq r - 1$. Damit erhalten wir eine Matrix A'', dass

$$a''_{1j_r} = \cdots = a''_{r-1,j_r} = 0.$$

Ist A''' die Matrix, die aus A'' durch Streichen der r-ten Zeile entsteht, so können wir nach der Induktionsvoraussetzung die Matrix A''' in eine Matrix in normierter Zeilenstufenform überführen. Zusammen mit der r-ten Zeile der Matrix A'' ergibt sich damit die gewünschte $m \times n$-Matrix B in normierten Zeilenstufenform. □

Bemerkung 3.11 Eine Matrix in normierter Zeilenstufenform nennt man oft **Gauss-Jordan** Matrix.

3.5 Lösungsmenge eines LGS

Definition 3.12 Sei (A, b) eine erweiterte Koeffizienten Matrix des LGS in Zeilenstufenform. Wir werden $x_{j_1}, \ldots, x_{j_r}$ **gebundene** Variablen nennen. Die Variablen x_i mit $i \notin \{j_1, \ldots, j_r\}$ werden **freie Variablen** genannt.

Bemerkung 3.13 Durch Umnumerierung der Variablen kann man immer erreichen, dass $x_{r+1}, \ldots, x_n$ genau die freien und $x_1, \ldots, x_r$ genau die gebundenen Variablen sind. Wir werden es oft annehmen, weil man damit mit Bezeichnungen sparen kann.

Satz 3.14 *Sei (A, b) eine Matrix in Zeilenstufenform. Dann gilt*
(1) $\text{Lös}(A, b) \neq \emptyset \Rightarrow r(A) = r(A, b)$.
(2) *Ist die Matrix (A, b) in normierter Zeilenstufenform, so dass $j_1 = 1, j_2 = 2, \ldots, j_r = r$ und $r(A) = r(A, b)$, so kann man die Lösungsmenge* $\text{Lös}(A, b)$ *in der folgenden Form aufschreiben:*

$$\text{Lös}(A, b) = (b_1 - \sum_{i=1}^{n-r} a_{1,r+i}\lambda_i, b_2 - \sum_{i=1}^{n-r} a_{2,r+i}\lambda_i, \ldots, b_r - \sum_{i=1}^{n-r} a_{r,r+i}\lambda_i, \lambda_1, \ldots, \lambda_{n-r}),$$

wobei $(\lambda_1, \lambda_2, \ldots, \lambda_{n-r})$ alle Vektoren aus $\mathbb{R}^{n-r}$ durchläuft.

Beweis. (1) Ist $r(A) \neq r(A, b)$, so gibt es eine Zeile in (A, b) der Form

$$(0, \ldots, 0, b_i), \quad b_i \neq 0.$$

Die zugehörige Gleichung $0 \cdot x_1 + \cdots + 0 \cdot x_n = b_i$ besitzt keine Lösung und damit $\text{Lös}(A, b) = \emptyset$.
(2) Jede Gleichung $\langle Z_k(A), x\rangle = b_k$ $(1 \leq k \leq r)$ hat die Form

$$x_k + \sum_{i=1}^{n-r} a_{i,r+i}x_{r+i} = b_k.$$

Wir können den $n-r$ freien Variablen $x_{r+1}, \ldots, x_n$ beliebige Werte $\lambda_1, \ldots, \lambda_{n-r}$ angeben. Damit sind die Werte der gebunden Variablen $x_1, \ldots, x_r$ eindeutig bestimmt. Also gilt

$$x_{r+1} = \lambda_1, \ x_{r+2} = \lambda_2, \ \ldots, x_n = \lambda_n$$

und

$$x_k = b_k - \sum_{i=1}^{n-r} a_{i,r+i}x_{r+i} = b_k - \sum_{i=1}^{n-r} a_{i,r+i}\lambda_i \ \ k = 1, \ldots, r.$$

□

Definition 3.15 Die oben gennante Schreibweise von $\text{Lös}(A, b)$ durch die reellen Zahlen $\lambda_1, \ldots, \lambda_r$ heißt die **Parametrisierung der allgemeinen Lösung** des LGS mit erweiterter Koeffizientenmatrix (A, b).

Korollar 3.16 *Sei (A, b) in Zeilenstufenform. Die Lösungsmenge* $\text{Lös}(A, b)$ *ist nicht leer genau dann, wenn $r(A) = r(A, b)$.*

Beweis. Aus dem Satz 3.14 (2) folgt, dass man immer eine Lösung finden kann, falls $r(A) = r(A, b)$. Deswegen gilt $r(A) = r(A, b) \Rightarrow \text{Lös}(A, b) \neq \emptyset$. Zusammen mit 3.14 (1) ergibt sich

$$r(A) = r(A, b) \Leftrightarrow \text{Lös}(A, b) \neq \emptyset.$$

□

Korollar 3.17 *Sei (A, b) in Zeilenstufenform. Die Lösungsmenge* $\text{Lös}(A, b)$ *besteht aus einem Element genau dann, wenn $r(A) = r(A, b) = n$. Ist die Matrix (A, b) in normierter Zeilenstufen form und $r(A) = r(A, b) = n$, so bilden die ersten n Komponenten von b die einzige Lösungs des LGS.*

Beweis. Aus 3.14 und 3.17 folgt, dass die Menge $\text{Lös}(A, b)$ genau dann aus einem einzigen Element besteht, wenn es keine freie Variabel gibt und $r(A) = r(A, b)$. Die letzte Bedingung ist erfühlt genau dann, wenn die Anzhal von Null verschiedenen Zeilen in A gleich n ist, d.h. $r(A) = n$. □

Satz 3.18 *Ist A eine beliebige $m \times n$-Matrix mit $m < n$, so besitzt das LGS mit der erweiterten Koeffizentenmatrix $(A, 0)$ eine von Null verschiedene Lösung.*

Beweis. Sei A' eine Matrix in Zeilenstufenform, die aus A durch endlich viele elementare Zeilnumformungen entstanden ist. Dann $r(A') \leq m < n$ (die Anzahl von Null verschiedenen Zeilen kann nur kleiner werden). Damit gibt es mindestens eine freie Variabel x_i des LGS mit der Koeffizientenmatrix $(A', 0)$. Wir setzen die Werte von freien Variablen gleich 1. Wegen 3.14 sind die Werte von gebundenen Variablen $x_{j_1}, \ldots, x_{j_r}$ eindeutig bestimmt. Dadurch erhalten wir eine eine von Null verschiedene Lösung (mindestens für eine Komponente x_i dieser Lösung gilt $x_i = 1 \neq 0$).

□

4 Matrizen

4.1 Operationen mit Matrizen

Mit $M(m \times n; \mathbb{R})$ bezeichnen wir die Menge aller $m \times n$-Matrizen mit reellen Einträgen. Die Elemente aus der Menge $M(n \times n; \mathbb{R})$ heißen **quadratische Matrizen**.

Definition 4.1 Seien $A, B \in M(m \times n; \mathbb{R})$. Die **Summe** $A + B$ zweier Matrizen ist die Matrix $C \in M(m \times n; \mathbb{R})$ mit Einträgen

$$c_{ij} := a_{ij} + b_{ij}, \;\; 1 \leq i \leq m, 1 \leq j \leq n.$$

Definition 4.2 Sei $A \in M(m \times n; \mathbb{R})$ und $\lambda \in \mathbb{R}$ eine reelle Zahl. Das **Produkt von A mit dem Skalar** $\lambda \in \mathbb{R}$ ist die Matrix $C \in M(m \times n; \mathbb{R})$ mit Einträgen

$$c_{ij} := \lambda \cdot a_{ij}, \;\; 1 \leq i \leq m, 1 \leq j \leq n.$$

Definition 4.3 Sei $A \in M(m \times n; \mathbb{R})$. Die Matrix $C \in M(n \times m; \mathbb{R})$ mit Einträgen

$$c_{ij} := a_{ji}$$

heißt **zu A transponierte Matrix** und wird mit ${}^tA$ bezeichnet.

Eine der wichtigsten Operationen mit Matrizen ist Multiplikation von Matrizen:

Definition 4.4 Seien $A \in M(m \times n; \mathbb{R})$ und $B \in M(n \times l; \mathbb{R})$. Das **Produkt** $A \cdot B$ (oder kurz AB) zweier Matrizen ist die Matrix $C \in M(m \times l; \mathbb{R})$ mit Einträgen

$$c_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + \cdots + a_{in} b_{nj}, \;\; 1 \leq i \leq m, 1 \leq j \leq l.$$

Der Eintrag c_{ij} von C entsteht als Skalarprodukt $\langle Z_i(A), S_j(B) \rangle$ der i-ten Zeile von A und der j-ten Spalte von B. (Beachten Sie, dass die Anzahl n gleich der Anzahl von Spalten in A und der Anzahl von Zeilen in B ist).

Bemerkung 4.5 Das Produkt von Matrizen hängt von Reihenordnung ab. Man kann nicht erwarten, dass für $n \geq 2$ gilt $A \cdot B = B \cdot A$, z. B.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \neq \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Bemerkung 4.6 Durch Multiplikation von Matrizen können wir ein LGS mit einer erweiterten Koeffizienten Matrix (A, b) kurz als

$$Ax = b$$

aufschreiben, wobei $A \in M(m \times n; \mathbb{R})$, und die $x \in M(n \times 1; \mathbb{R})$ und $bM(m \times 1; \mathbb{R})$ sind.

Satz 4.7 *Das Produkt von Matrizen ist assoziativ: sind A, B, C drei Matrizen mit $A \in M(p \times q; \mathbb{R})$, $B \in M(q \times r; \mathbb{R})$, $C \in M(r \times s; \mathbb{R})$, so ist*

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Beweis. Seien $F := (A \cdot B) \cdot C$ und $G := A \cdot (B \cdot C)$. Dann gilt

$$f_{ij} = \sum_{l=1}^{r} (\sum_{k=1}^{q} a_{ik} b_{kl}) c_{lj} = \sum_{l=1}^{r} \sum_{k=1}^{q} a_{ik} b_{kl} c_{lj} = \sum_{k=1}^{q} \sum_{l=1}^{r} a_{ik} b_{kl} c_{lj} = \sum_{k=1}^{q} a_{ik} (\sum_{l=1}^{r} b_{kl} c_{lj}) = g_{ij}.$$

□

Definition 4.8 Die quadratische $n \times n$-Matrix

$$E_n := \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

heißt **Einheitsmatrix**. Ist A eine quadratische Matrix und $k \in \mathbb{N}$, so definiert man

$$A^k := \underbrace{A \cdot A \cdots A}_{k}.$$

Es ist einfach nachzuweisen, dass für Matrizen die folgenden Rechenregeln gelten:

Proposition 4.9 *Sind $A, A' \in M(m \times n; \mathbb{R})$, $B, B' \in M(n \times l; \mathbb{R})$ und $\lambda \in \mathbb{R}$, so gilt*

(1) *Distributivgesetze:*

$$A \cdot (B + B') = A \cdot B + A \cdot B';$$

$$(A + A') \cdot B = A \cdot B + A' \cdot B;$$

(2) $A \cdot (\lambda B) = (\lambda A) B = \lambda (A \cdot B)$;
(3) ${}^t(A \cdot B) = {}^t B \cdot {}^t A$;
(4) $E_m \cdot A = A \cdot E_n = A$.

4.2 Elementare Matrizen

Definition 4.10 Die folgenden quadratischen $m \times m$-Matrizen I_{ij}, $E_{ij}(\lambda)$, $D_i(\lambda)$ werden **Elementarmatrizen** gennant:

(1) I_{ij} ist die $m \times m$-Matrix, die aus E_m durch Vertauschung der i-ten und j-ten Spalten entstanden ist;
(2) $E_{ij}(\lambda)$ ist die $m \times m$-Matrix, die aus E_m durch Addition der λ-fachen j-ten Zeile zur i-ten Zeile entstanden ist;
(3) $D_i(\lambda)$ ist die $m \times m$-Matrix, die aus E_m durch Multiplikation der i-ten Zeile mit Skalar $\lambda \in \mathbb{R}$ $(\lambda \neq 0)$ entstanden ist.

Proposition 4.11 *Sei A eine $m \times n$-Matrix. Die elementaren Zeilenumformungen von Typen $1, 2, 3$ können durch Multiplikation mit Elementarmatrizen wie folgt ausgedruckt werden:*

$$A \to A' := I_{ij}A \quad \text{Typ 1;}$$

$$A \to A' := E_{ij}(\lambda)A \quad \text{Typ 2;}$$

$$A \to A' := D_i(\lambda)A \quad \text{Typ 3.}$$

Man kann die elementaren Spaltenumformungen von Typen 1,2,3 genau wie die für Zeilen definieren. Damit erhalten wir:

Proposition 4.12 *Sei A eine $n \times m$-Matrix. Die elementaren Spaltenumformungen vom Typen $1, 2, 3$ können durch Multiplikation mit Elementarmatrizen wie folgt ausgedruckt werden:*

$$A \to A' := AI_{ij} \quad \text{Typ 1;}$$

$$A \to A' := AE_{ij}(\lambda) \quad \text{Typ 2;}$$

$$A \to A' := AD_i(\lambda) \quad \text{Typ 3.}$$

Satz 4.13 *Jede Matrix A läßt sich mit elementaren Spalten- und Zeilenumformungen in einer Matrix der Gestalt*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

überfuhren.

4.3 Einfachste lineare Matrizengleichungen

Definition 4.14 Seien A eine $m \times n$-Matrix, B eine $m \times l$-Matrix und C eine $l \times n$-Matrix. Die Gleichungen

$$A \cdot X = B,$$

$$Y \cdot A = C,$$

wobei $X \in M(n \times l; \mathbb{R})$ und $Y \in M(l \times m; \mathbb{R})$ unbekannte Matrizen sind, werden **einfachste lineare Matrizengleichungen** genannt.

Wegen

$$Y \cdot A = C \Leftrightarrow A^t \cdot Y^t = C^t$$

genügt es, nur die Gleichungen der Form $A \cdot X = B$ zu betrachten. Wir bemerken, dass man $A \cdot X = B$ durch die folgenden l LGS für Spalten $S_1(X), \ldots, S_l(X)$ interpretieren kann:

$$A \cdot S_1(X) = S_1(B),$$

$$A \cdot S_2(X) = S_2(B),$$

$$\ldots$$

$$A \cdot S_l(X) = S_l(B).$$

Die erweiterten Matrizen von l Systemen haben Gestalt

$$(A, S_1(B)), \quad (A, S_2(B)), \quad \ldots, \quad (A, S_l(B)).$$

Es ist einfacher diese Systeme gleichzeitig zu lösen und das Gaußsche Eliminationsverfahren direkt auf die Matrix

$$(A, S_1(B), S_2(B), \ldots, S_l(B)) = (A, B)$$

anzuwenden.

Proposition 4.15 *Sei A eine quadratische $n \times n$-Matrix. Ist $B \in M(n \times n; \mathbb{R})$ eine Lösung der Matrizengleichung $AX = E_n$ und $B' \in M(n \times n; \mathbb{R})$ eine Lösung der Matrizengleichung $XA = E_n$, so ist $B = B'$.*

Beweis.

$$B' = B' \cdot E_n = B' \cdot (A \cdot B) = (B' \cdot A) \cdot B = E_n \cdot B = B.$$

□

Definition 4.16 Sei A eine quadratische $n \times n$-Matrix. Eine Matrix B mit $AB = BA = E_n$ heißt **die zu A inverse Matrix** und wird mit A^{-1} bezeichnet (falls solche Matrix B existiert).

Proposition 4.17 *Seien A und B zwei quadratischen $n \times n$-Matrizen, so dass die inversen Matrizen A und B existieren. Dann $B^{-1}A^{-1}$ ist die inverse Matrix zum Produkt AB.*

Beweis.

$$(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = B^{-1} \cdot ((A^{-1} \cdot A) \cdot B) = B^{-1} \cdot E_n \cdot B = B^{-1} \cdot B = E_n.$$

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1}) = A \cdot E_n \cdot A^{-1} = A \cdot A^{-1} = E_n.$$

□

Beispiel 4.18 Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Für die Bestimmung der inversen Matrix A^{-1} betrachten man die erweiterte Matrix (A, E_2):

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{pmatrix}.$$

Die inverse Matrix ist $A^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$.

Bemerkung 4.19 Es ist einfach zu zeigen, dass die inverse Matrizen zu Elementarmatrizen wieder Elementarmatrizen sind:

$$I_{ij}^{-1} = I_{ij}, \quad E_{ij}(\lambda)^{-1} = E_{ij}(-\lambda), \quad D_i(\lambda)^{-1} = D_i(\lambda^{-1}).$$

5 Unterräume, Basen und Dimension

5.1 Linearkombination, lineare Unbhängigkeit

Definition 5.1 Sei $\{v_1, \dots, v_k\}$ eine Menge von Vektoren in $\mathbb{R}^n$. Man sagt, dass ein Vektor $x \in \mathbb{R}^n$ als **Linearkombination von** $v_1, \dots, v_k$ **darstellbar**, wenn gilt

$$x = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

mit einigen reellen Zahlen $\lambda_1, \dots, \dots_k$. Die reellen Zahlen $\lambda_1, \dots, \dots, \lambda_k$ nennt man **Koeffizienten der Linearkombination**. Sind alle $\lambda_1, \dots, \dots, \lambda_k$ gleich 0, so nennt man $\lambda_1 v_1 + \cdots + \lambda_k v_k$ **triviale Linearkombination**.

Definition 5.2 Sei $\{v_1, \dots, v_k\}$ eine Menge von Vektoren in $\mathbb{R}^n$. Mit

$$\mathrm{Span}(v_1, \dots, v_k)$$

bezeichnen wir die Menge aller Vektoren, die als Linearkombination von $v_1, \dots, v_k$ darstellbar sind. Die Menge $\mathrm{Span}(v_1, \dots, v_k)$ heißt **lineare Hülle** von Vektoren $v_1, \dots, v_k$.

Definition 5.3 Die Vektoren $v_1, \dots, v_k$ heißen **linear unabhängig**, wenn Nullvektor nur als triviale Linearkombination von $v_1, \dots, v_k$ darstellbar ist, d.h.,

$$0 = \lambda_1 v_1 + \cdots + \lambda_k v_k \ \Rightarrow \lambda_1 = \cdots = \lambda_k = 0.$$

Sind Vektoren $v_1, \dots, v_k$ nicht linear unabhängig, so heißen $v_1, \dots, v_k$ **linear abhängig**. Die lineare Abhängigkeit von $v_1, \dots, v_k$ bedeutet, dass es eine nichttriviale Linearekombination $\lambda_1 v_1 + \cdots + \lambda_k v_k$ gibt (mindestens für eine Zahl λ_i gilt $\lambda_i \neq 0$), die gleich Nullvektor ist.

Proposition 5.4 *Ein System ist linear abhängig genau dann, wenn ein Vektor $v_i \in S$ gibt, die als Linearkombination von $S \setminus \{v_i\}$ darstellbar ist.*

5.2 Untervektorräume und Basen

Definition 5.5 Eine nichtleere Teilmenge $L \subset \mathbb{R}^n$ heißt **Untervektorraum**, wenn die folgenden Bedingungen erfühlt sind:
(1) $x, y \in L \ \Rightarrow \ x + y \in L$;
(2) $x \in L, \ \lambda \in \mathbb{R} \Rightarrow \lambda x \in L$.

Proposition 5.6 *Sei $\{v_1, \dots, v_k\}$ eine Menge von Vektoren in $\mathbb{R}^n$. Dann ist* $\mathrm{Span}(v_1, \dots, v_k)$ *ein Untervektorraum.*

Beweis Wir müssen zeigen, dass für $L = \text{Span}(v_1, \ldots, v_k)$ die Bedingungen (1) und (2) in 5.5 erfüllt sind:
Seien $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in L$. Dann gibt es $\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_k \in \mathbb{R}$, so dass

$$x = \lambda_1 v_1 + \cdots + \lambda_k v_k,$$

$$y = \mu_1 v_1 + \cdots + \mu_k v_k.$$

Wir erhalten $x + y \in L$, weil der Vektor $x + y$ zur Linearkombination $(\lambda_1 + \mu_1)v_1 + \cdots + (\lambda_k + \mu_k)v_k$ gleich ist. Ferner für alle $\Lambda \in \mathbb{R}$ gilt $\lambda x \in L$, weil λx zur Linearkombination $\lambda\lambda_1 v_1 + \cdots + \lambda\lambda_k v_k$ gleich ist.

□

Definition 5.7 Sei $L \subset \mathbb{R}^n$ ein Untervektorraum und $\{v_1, \ldots, v_d\} \subset L$ eine Teilmenge. Die Menge $\{v_1, \ldots, v_d\}$ heißt eine **Basis von** L, wenn gilt
(1) $v_1, \ldots, v_d$ sind linear unabhängig;
(2) jeder Vektor $x \in L$ ist als Linearkombination von $v_1, \ldots, v_d$ darstellbar.

Satz 5.8 *Jeder Vektorunterraum $L \subset \mathbb{R}^n$ besitzt eine Basis.*

Beweis. Ist $L = \{0\}$, so kann man als Basis die leere Menge von Vektoren nehmen. Sei nun $L \neq \{0\}$. Dann gibt es einen Vektor $v \in L$ mit $v \neq 0$. Das System $\{v\}$ ist linear unabhängig: $\lambda v = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n) = 0 \Leftrightarrow \lambda = 0$. Sei $\{v_1, \ldots, v_d\} \subset L$ die lineare unabhähgige Teilmenge in L, die maximale Anzahl $d \geq 1$ der Vektoren besitzt. Wir schreiben die Komponenten $(a_{1i}, \ldots, a_{ni})$ von v_i als i-te Spalte $S_i(A)$ einer $n \times d$-Matrix A auf. Aus dem 3.18 folgt, dass d darf nicht größer als n sein (sonst wären die Spalten der Matrix A linear abhängig).
Wir zeigen, dass $\{v_1, \ldots, v_d\}$ eine Basis von L ist. In der Tat, die erste Bedingung in 5.7 ist offensichtlich erfüllt. Ist $x \in L$ ein beliebiger Vektor, so ist die Menge $\{v_1, \ldots, v_d, x\}$ linear abhängig. Damit gibt es eine nichttriviale Linearkombination

$$0 = \lambda_1 v_1 + \cdots + \lambda_d v_d + \lambda_{d+1} x.$$

Wir bemerken, dass λ_{d+1} nicht gleich 0 seien darf (sonst wären $v_1, \ldots, v_d$ linear abhängig). Nun ist x als Linearkombination

$$x = (-\lambda_1/\lambda_{d+1})v_1 + \cdots + (-\lambda_d/\lambda_{d+1})v_d$$

darstellbar. Damit gilt (2) aus 5.7.

□

Satz 5.9 *Seien $\{v_1, \ldots, v_k\} \subset L$ und $\{u_1, \ldots, u_l\} \subset L$ zwei Basen des Unterraums L. Dann gilt $k = l$.*

Beweis. Seien $\{v_1, \ldots, v_k\}, \{u_1, \ldots, u_l\} \subset L$ zwei Basen. Es genügt zu zeigen, dass z. B. $k > l$ nicht möglich ist.
Sei $k > l$. Dann schreiben wir jeden Vektor v_j als Linearkombination

$$v_j = a_{1j}u_1 + a_{2j}u_2 + \cdots + a_{lj}u_l$$

auf. Wir betrachten die Koeffizienten $(a_{1j}, \ldots, a_{lj})$ als j-te Spalte $S_j(A)$ einer $l \times k$-Matrix A. Aus dem Satz 3.18 folgt, dass die Zeilen der Matrix A linear abhängig sind, d. h., es existiert eine Lösung $0 \neq \alpha := (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$ des LGS mit $({}^tA, 0)$. Daraus folgt

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0.$$

Widerspruch.

□.

Definition 5.10 Sei $L \subset \mathbb{R}^n$ ein Vektorunteraum. Die Anzahl der Vektoren in einer Basis von L nennt man **Dimension von** L.

Proposition 5.11 *Sind $L_1, \ldots, L_m \subset \mathbb{R}^n$ Vektorunterräume, so ist $L_1 \cap \cdots \cap L_m \subset \mathbb{R}^n$ ein Vektorunterraum.*

5.3 Lösungsmengen und Untervektorräume

Definition 5.12 Ein LGS mit erweiterter Koeffizientenmatrix (A, b) heißt **homogen**, wenn gilt $b = 0$ (d.h. b ist eine Null-Spalte).

Satz 5.13 *Sei A eine $m \times n$-Matrix. Die Lösungsmenge eines homogenen Gleichungssystems* $\text{Lös}(A, 0) \subset \mathbb{R}^n$ *ist ein Vektorunterraum.*
Ist A eine $m \times n$-Matrix in Zeilenstufenform, so ist dim$\text{Lös}(A, 0) = n - r$*, wobei $r = r(A)$ die Anzahl von Null verschiedenen Zeilen in A ist.*

Beweis. Wir schreiben das homogene LGS kurz als

$$\langle Z_i(A), x\rangle = 0, \quad i = 1, 2, \ldots, m$$

auf. Sind $x, y \in \text{Lös}(A, 0)$ zwei beliebige Lösungen, so ist $x + y \in \text{Lös}(A, 0)$, weil für alle i gilt

$$\langle Z_i(A), x + y\rangle = \langle Z_i(A), x\rangle + \langle Z_i(A), y\rangle = 0 + 0 = 0, \quad i = 1, 2, \ldots, m.$$

Ferner erhalten wir für jede Skalar $\lambda \in \mathbb{R}$:

$$\langle Z_i(A), \lambda x\rangle = \lambda\langle Z_i(A), x\rangle = 0, \;\; i = 1, 2, \ldots, m.$$

Also ist $\text{Lös}(A, 0) \subset \mathbb{R}^n$ ist ein Vektorunterraum.
Sei nun A eine $m \times n$-Matrix in Zeilenstufenform und $r := r(A)$ die Anzahl von Null verschiedenen Zeilen in A. Um $\text{Lös}(A, 0) = n - r$ zu zeigen, werden wir eine Basis von $\text{Lös}(A, 0)$ finden, die genau aus $n - r$ Vektoren $v_1, \ldots, v_{n-r}$ besteht. Sei A' eine $m \times n$-Matrix in normierter Zeilenstrufenform, die aus A durch elementare Zeilenumformungen vom Typ 2 und 3 entstanden ist (s. 3.10). Wegen $r = r(A') = r(A)$ und $\text{Lös}(A, 0) = \text{Lös}(A', 0)$ genügt es eine o.g. Basis für den Fall $A = A'$ zu finden. Also können wir ohne Beschränkung der Allgemeinheit annehmen, dass A selbe in *normierter* Zeilenstrufenform ist. Unter diesen Umständen dürfen wir die Parametrisierung aus 3.14 (2) benutzen. Also ist die Menge $\text{Lös}(A, 0)$ durch $n - r$ Skalare $\lambda_1, \ldots, \lambda_{n-r}$ wie folgt parametrisiert:

$$\text{Lös}(A, 0) = (-\sum_{i=1}^{n-r} a_{1,r+i}\lambda_i, -\sum_{i=1}^{n-r} a_{2,r+i}\lambda_i, \ldots, -\sum_{i=1}^{n-r} a_{r,r+i}\lambda_i, \lambda_1, \ldots, \lambda_{n-r}).$$

Wir definieren

$$v_1 := (\underbrace{-a_{1,r+1}, -a_{2,r+1}, \cdots, -a_{r,r+1}}_{r}, \underbrace{1, 0, \cdots, 0}_{n-r});$$

$$v_2 := (\underbrace{-a_{1,r+2}, -a_{2,r+2}, \cdots, -a_{r,r+2}}_{r}, \underbrace{0, 1, \cdots, 0}_{n-r});$$

$$\cdots$$

$$v_{n-r} := (\underbrace{-a_{1\,n}, -a_{2\,n}, \cdots, -a_{r\,n}}_{r}, \underbrace{0, 0, \cdots, 1}_{n-r}).$$

Aus der Parametrisierung folgt, dass jeder Vektor $x \in \text{Lös}(A, 0)$ eine Darstellung als Linearkombination $\lambda_1 v_1 + \cdots + \lambda_{n-r} v_{n-r}$ besitz. Andererseits sind die Vektoren $v_1, \ldots, v_{n-r}$ linear unabhängig: ist $v = \lambda_1 v_1 + \cdots + \lambda_{n-r} v_{n-r}$ ein Nullvektor, so sind die letzten $n - r$ Komponenten von

$$v = (\underbrace{*, \cdots, *}_{r}, \underbrace{\lambda_1, \ldots, \lambda_{n-r}}_{n-r})$$

gleich Null, d.h. $\lambda_1 = \cdots = \lambda_{n-r} = 0$. Damit haben wir gezeigt, dass Vektoren $v_1, \ldots, v_{n-r}$ eine Basis von $\text{Lös}(A, 0)$ bilden.

□

Satz 5.14 *Sei $x = (x_1, \ldots, x_n) \in \text{Lös}(A, b)$ eine beliebige Lösung des LGS mit der erweiterten Koeffizientenmatrix (A, b). Dann gilt*

$$\text{Lös}(A, b) = x + \text{Lös}(A, 0),$$

wobei die Menge $x + \text{Lös}(A, 0)$ aus aller Summen $\{x + y : y \in \text{Lös}(A, 0)\}$ besteht.

Beweis. Sei $y \in \text{Lös}(A, 0)$ eine beliebige Lösung des homogenen GLS und x ist ein beliebiges Element von $\text{Lös}(A, b)$. Dann gilt

$$\langle Z_i(A), y\rangle = 0, \quad i = 1, 2, \ldots, m$$

und

$$\langle Z_i(A), x\rangle = b_i, \quad i = 1, 2, \ldots, m.$$

Für den Vektor $x + y$ erhalten wir

$$\langle Z_i(A), x + y\rangle = \langle Z_i(A), x\rangle + \langle Z_i(A), y\rangle = b_i + 0 = b_i, \quad i = 1, 2, \ldots, m.$$

Also ist $x + y$ eine Lösung aus $\text{Lös}(A, b)$ und damit gilt die Inklusion

$$\text{Lös}(A, b) \supset x + \text{Lös}(A, 0).$$

Ist x' noch ein beliebiges Element von $\text{Lös}(A, b)$, so ist die Differenz $x' - x$ ein Element von $\text{Lös}(A, 0)$:

$$\langle Z_i(A), x' - x\rangle = \langle Z_i(A), x'\rangle - \langle Z_i(A), x\rangle = b_i - b_i = 0, \quad i = 1, 2, \ldots, m.$$

Deswegen läßt sich jedes $x' \in \text{Lös}(A, b)$ als Summe $x + y$ darstellen, wobei $y := x' - x \in\in \text{Lös}(A, 0)$. Damit gilt die Inklusion

$$\text{Lös}(A, b) \subset x + \text{Lös}(A, 0).$$

□

Bemerkung 5.15 Aus dem Satz 5.14 folgt, dass die Lösungsmenge $\text{Lös}(A, b) = x + \text{Lös}(A, 0)$ sich als mit Vektor x "verschobenen" Vektorrunterraum $\text{Lös}(A, 0)$ darstellen läßt, falls $\text{Lös}(A, b)$ nicht leer ist, d.h., es mindestens ein Element $x \in \text{Lös}(A, b)$ gibt.

5.4 Rang einer Matrix

Definition 5.16 Sei A eine $m \times n$-Matrix. Die Dimension

$$ZR(A) := \dim \operatorname{Span}(Z_1(A), \ldots, Z_m(A))$$

nennt man **Zeilenrang von** A. Die Dimension

$$SR(A) := \dim \operatorname{Span}(S_1(A), \ldots, S_n(A))$$

wird **Spaltenrang von** A genannt.

Beispiel 5.17 Sei

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 0 & 6 \end{pmatrix}.$$

Dann gilt

$$\operatorname{Span}(Z_1(A), Z_2(A)) = \{\lambda(1, 0, 3) \ : \ \lambda \in \mathbb{R}\}$$

und

$$\operatorname{Span}(S_1(A), S_2(A), S_3(A)) = \left\{ \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} \ : \ \lambda \in \mathbb{R} \right\}.$$

Deswegen gilt

$$ZR(A) = SR(A) = 1.$$

Wir benötigen die folgende Aussage:

Satz 5.18 *Der Zeilenrang $ZR(A)$ ist zum Spaltenrang $SR(A)$ gleich.*

Für den Beweis brauchen wir zwei Hilfsaussagen 5.19 und 5.20.

Lemma 5.19 *Sei A' aus A durch endlich viele elementaren Zeilenumformungen entstandene Matrix. Dann gilt*

$$\operatorname{Span}(Z_1(A), \ldots, Z_m(A)) = \operatorname{Span}(Z_1(A'), \ldots, Z_m(A'))$$

und damit $ZR(A) = ZR(A')$.

Beweis. Es genügt nur eine Zeilenumformung von A zu betrachten. Bei dieser Zeilenumformung läßt sich jede Zeile von A' als Linearkombination der Zeilen von A darstellen. Durch inverse Zeilenumformung (s. 4.19) erhalten wir, dass auch jede Zeile von A sich als Linearkombination der Zeilen von A' darstellen läßt. Deswegen sind die Vektorunterräume $\operatorname{Span}(Z_1(A), \ldots, Z_m(A))$ und $\operatorname{Span}(Z_1(A'), \ldots, Z_m(A'))$ gleich. □

Lemma 5.20 *Sei A' aus A durch endlich viele elementaren Zeilenumformungen entstandene Matrix und $1 \leq j_1 < j_2 < \cdots j_k \leq n$. Dann bilden die Spalten $S_{j_1}(A), \ldots, S_{j_k}(A)$ eine Basis von $\mathrm{Span}(S_1(A), \ldots, S_n(A))$ genau dann, wenn die Spalten $S_{j_1}(A'), \ldots, S_{j_k}(A')$ eine Basis von $\mathrm{Span}(S_1(A'), \ldots, S_n(A'))$ bilden. Insbesondere gilt $SR(A) = SR(A')$.*

Beweis. Hauptidee: wir verwenden die Gleichung $\mathrm{Lös}(A, 0) = \mathrm{Lös}(A', 0)$.
Ohne Einschränkung der Allgemeinheit können wir annehmen, dass $j_1 = 1, \ldots, j_k = k$ ist. Nun zeigen wir, dass die beiden Eigenschaften der Basis für $S_1(A), \ldots, S_k(A)$ und $S_1(A'), \ldots, S_k(A')$ äquivalent sind.
(1) Wegen der Gleichung

$$\lambda_1 S_1(A) + \cdots + \lambda_k S_k(A) = A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

sind Spalten $S_1(A), \ldots, S_k(A)$ linear unabhängig genau dann, wenn jede Lösung $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathrm{Lös}(A, 0)$ des homogenen LGS $Ax = 0$ trivial ist, falls $\lambda_{k+1} = \cdots = \lambda_n = 0$ ist. Wegen $\mathrm{Lös}(A, 0) = \mathrm{Lös}(A', 0)$ (s. 3.6) ist die lineare Unabhängigkeit von $S_1(A), \ldots, S_k(A)$ zur linearen Unabhängigkeit von $S_1(A'), \ldots, S_k(A')$ äquivalent.
(2) Eine Spalte $S_j(A)$ $(j > k)$ ist als Linearkombination von $S_1(A), \ldots, S_k(A)$ darstellbar genau dann, wenn der Vektor $(\lambda_1, \ldots, \lambda_k, 0, \ldots, \underbrace{-1}_{j}, \ldots, 0)$ eine Lösung von LGS $Ax = 0$ ist. Wegen $\mathrm{Lös}(A, 0) = \mathrm{Lös}(A', 0)$ (s. 3.6) ist die Darstellbarkeit von $S_j(A)$ als Linearkombination von $S_1(A), \ldots, S_k(A)$ zur Darstellbarkeit von $S_j(A')$ als Linearkombination von $S_1(A'), \ldots, S_k(A')$ äquivalent.

□

Beweis von 5.18. Aus 5.19 und 5.20 folgt, dass sich $ZR(A)$ und $SR(A)$ bei elementaren Zeilenoperationen. Mit gleichen Überlegungen kann man zeigen, dass sich $ZR(A)$ und $SR(A)$ bei elementaren Spaltenoperationen $ZR(A)$ und $SR(A)$ nicht ändern. Wegen 4.13 genügt es die Aussage nur für die Matrix der Form $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ zu beweisen, die in diesem Fall offensichtlich ist:

$$ZR \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = SR \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r.$$

□

Definition 5.21 Der gemeinsame Wert von $ZR(A)$ und $SR(A)$ einer Matrix A heißt **Rang der Matrix** A und wird mit Rang(A) (oder $R(A)$ bezeichnet.

6 Abbildungen

6.1 Abbildungen und ihre Eigenschaften

Definition 6.1 Sei $f : X \to Y$ eine Abbildung. Dann heißt f
(1) **injektiv**, wenn für beliebige $x_1, x_2 \in X$ gilt $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ (d.h. die Bilder verschiedener Elemente aus X müssen wieder voneinander verschieden sein.)
(2) **surjektiv**, wenn zu jedem $y \in Y$ es ein $x \in X$ gibt, so dass $f(x) = y$ (d.h. das Bild $f(X)$ von X ist gleich Y).
(iii) **bijektiv**, wenn f gleichzeitig surjektiv und injektiv ist.

Definition 6.2 Seien $f : X \to Y$, $g : Y \to Z$ Abbildungen. Dann bezeichnen wir mit $g \circ f$ die Abbildung $X \to Z$, welche das Element $x \in X$ in das Element $g(f(x)) \in Z$ überführt. Die Abbildung $g \circ f$ nennen wir die **Komposition** der Abbildungen f und g. Zwei Abblidungen $f_1 : X \to Y$ und $f_2 : X \to Y$ heißen **gleich**, wenn für alle $x \in X$ gilt $f_1(x) = f_2(x)$.

Proposition 6.3 *Seien*

$$f : X \to Y, \quad g : Y \to Z, \quad h : Z \to W$$

Abbildungen. Dann gilt das Assoziativgesetz:

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Beweis. Für jedes $x \in X$ gilt:

$$\begin{aligned} (h \circ (g \circ f))(x) &= h\,(g \circ f)(x)) \\ &= h\,(g(f(x))) \\ &= (h \circ g)\,(f(x)) = ((h \circ g) \circ f)\,(x). \end{aligned}$$

□

Definition 6.4 Wir bezeichnen mit id_X die **identische Abbildung** von X, d.h. die Abbildung von X in sich, die durch $id_X(x) = x$ für alle $x \in X$ definiert ist.

Definition 6.5 Sei $f : X \to Y$ eine Abbildung. Die Abbildung $g : Y \to X$ heißt die **Umkehrabbildung** von f, wenn gilt

$$g \circ f = id_X$$

und

$$f \circ g = id_Y.$$

Die Umkehrabbildung wird mit f^{-1} bezeichnet.

Proposition 6.6 *Sei $f : X \to Y$ eine Abbildung. Die Umkehrabbildung von f existiert genau dann, wenn f bijektiv ist.*

Beweis. Sei g die Umkehrabbildung. Aus $f \circ g = id_Y$ folgt $y = f(g(y))$ für alle $y \in Y$, d.h. f ist sujektiv. Aus $g \circ f = id_X$ folgt

$$f(x_1) = f(x_2) \Rightarrow g(f(x_1)) = g(f(x_2)) \text{ d.h. } x_1 = x_2.$$

Damit ist f injektiv.
Sei f bijektiv. Dann existiert für jedes $y \in Y$ ein einziges $x \in X$ mit $f(x) = y$. Wir definieren die Abbildung $g : Y \to X$ durch $g(y) := x$. Dann gilt $g \circ f = id_X$ und $f \circ g = id_Y$.

□

6.2 Lineare Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$

Definition 6.7 Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt **linear**, wenn die folgenden Bedingungen erfüllt sind:
(1) $\forall x, y \in \mathbb{R}^n$ gilt $f(x + y) = f(x) + f(y)$;
(2) $\forall x \in \mathbb{R}^n$ und $\forall \lambda \in \mathbb{R}$ gilt $f(\lambda x) = \lambda f(x)$.

Beispiel 6.8 Ist $f : \mathbb{R}^n \to \mathbb{R}^m$ eine Abbildung mit $f(x) = 0 \quad \forall x \in \mathbb{R}^n$, so heißt f **Nullabbildung**. Offensichtlich ist Nullabbildung linear (und wird mit 0 bezeichnet).

Beispiel 6.9 Ist $n = m$ und $f : \mathbb{R}^n \to \mathbb{R}^n$ die **identische Abbildung**: $f(x) = x \ \forall x \in \mathbb{R}^n$, so ist f auch linear (und wird mit id bzeichnet).

Es leicht zu zeigen:

Proposition 6.10 *Sei $A \in M(m \times n; \mathbb{R})$ eine beliebige Matrix. Mit φ_A bezeichnen wir die Abbildung $\mathbb{R}^n \to \mathbb{R}^m$:*

$$x \mapsto Ax,$$

wobei der Vektor $x \in \mathbb{R}^n$ als Matrix aus $M(n \times 1; \mathbb{R})$ aufgefasst wird. Dann ist φ_A eine lineare Abbildung.

Beweis. Wir verwenden die Eingenschaften der Matrizenmultiplikation:

$$A(x+y) = Ax + Ay \ \ \forall x, y \in \mathbb{R}^n,$$

$$A(\lambda x) = \lambda Ax \ \ \forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}.$$

□

Satz 6.11 *(Bijektion zwischen den Matrizen und linearen Abbildungen)*
Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ ist zu einer linearen Abbildung φ_A gleich, wobei A eine Matrix aus $M(m \times n; \mathbb{R})$ ist. Ferner gilt

$$\varphi_A = \varphi_B \Leftrightarrow A = B.$$

Damit entsteht eine Bijektion zwischen der Menge aller $m \times n$-Matrizen und der Menge aller linearen Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$.

Beweis. Es sei $e_1, \ldots, e_n$ die Standartbasis von $\mathbb{R}^n$. Ist $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung, so bezeichnen wir mit A die Matrix mit n Spalten

$$f(e_1), \ldots, f(e_n) \in \mathbb{R}^m,$$

d.h. $S_i(A) := f(e_i), \ \ i = 1, \ldots, n$. Sei $x = (x_1, \ldots, x_n)$ ein beliebiger Vektor aus $\mathbb{R}^n$. Wir schreiben $x = x_1 e_1 + \cdots + x_n e_n$. Da f linear ist, erhalten wir

$$f(x) = f(x_1 e_1 + \cdots + x_n e_n) = x_1 f(e_1) + \cdots + x_n f(e_n) = \sum_{i=1}^{n} x_i S_i(A) =$$

$$= A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Also gilt $f = \varphi_A$.
Ist $\varphi_A = \varphi_B$, so gilt $S_i(A) = \varphi_A(e_i) = \varphi_B = S_i(B)$ für alle $i = 1, \ldots, n$. Damit erhalten wir $A = B$. □

Proposition 6.12 *(Matrizenmultiplikation = Komposition linearer Abbildungen)*
Sind $B \in M(m \times n; \mathbb{R})$ und $A \in M(k \times m; \mathbb{R})$ beliebige Matrizen, so gilt

$$\varphi_{AB} = \varphi_A \circ \varphi_B.$$

Beweis. Ist $x \in \mathbb{R}^n$, so ist $y = \varphi_B(x) = Bx$. Ferner gilt $\varphi_A(y) = Ay$. Zusammen ergibt sich

$$(\varphi_A \circ \varphi_B)(x) = \varphi_A(\varphi_B(x)) = A(Bx) = (AB)x = \varphi_{AB}(x).$$

□

Definition 6.13 Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Die Menge

$$Ker\, f := \{x \in \mathbb{R}^n : f(x) = 0\}$$

heißt der **Kern** von f. Die Menge

$$Im\, f := \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ mit } f(x) = y\}$$

heißt das **Bild** von f.

Proposition 6.14 *Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Dann sind $Ker\, f \subset \mathbb{R}^n$ und $Im\, f \subset \mathbb{R}^m$ Vektorunterräume. Ist $f = \varphi_A$, so gilt*

$$Ker\, f = \text{Lös}(A, 0),\ Im\, f = \text{Span}(S_1(A), \ldots, S_m(A)).$$

und

$$dim\, Ker\, f + dim\, Im\, f = n.$$

6.3 Permutationen

Definition 6.15 Die bijektiven Abbildungen der Menge X auf sich werden **Permutationen von** X genannt. Die Menge aller Permutationen von X bezeichnen wir mit $S(X)$.

Definition 6.16 X bestehe aus n Elementen. Dann schreibt man S_n für $S(X)$. Sind $x_1, \ldots, x_n$ Elemente der Menge X, so bezeichnet man mit

$$\begin{bmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{bmatrix}$$

die Abbildung $\varphi : X \to X$

$$x_k \mapsto x_{i_k} \ \ (k = 1, \ldots, n))$$

(Damit identifizieren wir $X = \{x_1, x_2, \ldots, x_n\}$ mit der Menge $\{1, 2, \ldots, n\}$). Die Komposition von Permutationen σ und σ' wird mit $\sigma \cdot \sigma'$ (oder mit $\sigma\sigma'$) bezeichnet.

Ein Element $\sigma \in \mathcal{S}_n$ (Permutation) läßt sich durch die folgenden verschiedenen Formen darstellen:

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ l_1 & l_2 & l_3 & \cdots & l_n \end{bmatrix}, \quad \sigma(k) = l_k \ (k = 1, \ldots, n) \tag{1}$$

oder

$$\sigma = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}, \quad \sigma(i_k) = j_k \ (k = 1, \ldots, n) \tag{2}$$

Beispiel 6.17 .

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 & 5 & 4 \\ 4 & 5 & 3 & 1 & 2 \end{bmatrix}.$$

Definition 6.18 Sei

$$\sigma = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$$

eine Permutation Die Zahl

$$\prod_{l<m} \frac{(j_m - j_l)}{(i_m - i_l)}$$

heißt das **Signum** (oder **Vorzeichen**) von Permutation σ und wird mit $sgn(\sigma)$ bezeichnet. Es ist einfach zu zeigen, dass $sgn(\sigma) \in \{1, -1\}$. Eine Permutation $\sigma \in S_n$ heißt **gerade** (bzw. **ungerade**), wenn gilt $sgn(\sigma) = 1$ (bzw. $sgn(\sigma) = -1$).

Definition 6.19 Sei die Folge $f = \{k_1, k_2, \ldots, k_n\}$ eine Umordnung der Zahlen $1, 2, \ldots, n$. Ein Paar (i, j) mit $1 \leq i, j \leq n$ heißt ein **Fehlstand** (oder eine **Inversion**) von f, wenn $i < j$ und $k_i > k_j$. Die Anzahl der Fehlstände von f wird mit $z(f)$ bezeichnet. Aus 6.18 folgt: Ist

$$\sigma = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$$

eine Permutation und

$$f_1 := \{i_1, i_2, \ldots, i_n\}, \quad f_2 := \{j_1, j_2, \ldots, j_n\},$$

so ist

$$sgn(\sigma) = (-1)^{z(f_1)+z(f_2)}.$$

Lemma 6.20 *Sei*

$$f = \{k_1, k_2, \ldots, k_i, \ldots, k_j, \ldots, k_n\}$$

eine Umordnung der Zahlen $1, 2, \ldots, n$. *Ist die Folge*

$$f' = \{k'_1, k'_2, \ldots, k'_i, \ldots, k'_j, \ldots, k'_n\} := \{k_1, k_2, \ldots, k_j, \ldots, k_i, \ldots, k_n\}$$

durch die Vertauschung von k_i *und* k_j *aus* f *entstanden, so ist*

$$(-1)^{z(f')} = -(-1)^{z(f)}.$$

Beweis. 1. Fall: $j = i+1$. Ist $(l, m) \neq (i, j)$ ein Paar $1 \leq l < m \leq n$, so ist (l, m) ein Fehlstand von f genau dann, wenn (l, m) ein Fehlstand von f'. Andererseits gilt: Ist (i, j) ein Fehlstand von f, so ist (i, j) kein Fehlstand von f'. Ist (i, j) kein Fehlstand von f, so ist (i, j) ein Fehlstand von f'. Also gilt $z(f') = z(f)-1$ oder $z(f') = z(f)+1$ und damit $(-1)^{z(f')} = -(-1)^{z(f)}$. Nun betrachten wir den algemeinen Fall $j > i$ und bemerken, dass man die Vertauschung von k_i und k_j durch Vertauschungen von k_i und k_{i+1}, k_i und k_{i+2}, …, k_i und k_{j-1}, k_i und k_j und danach mit durch Vertauschungen von k_j und k_{j-1}, k_j und k_{j-2}, …, k_j und k_{i+1}. Insgesammt braucht man genau $2(j-i)-1$ Vertauschungen vom Typ, der schon im ersten Fall untersucht wurde. Deswegen gilt

$$(-1)^{z(f')} = (-1)^{2(j-i)-1}(-1)^{z(f)} = -(-1)^{z(f)}.$$

□

Satz 6.21 *Die Zahl* $sgn(\sigma)$ *hängt nicht von der Schreiweise von* σ *ab.*

Beweis. Seien

$$\sigma = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$$

und

$$\sigma = \begin{bmatrix} i'_1 & i'_2 & i'_3 & \cdots & i'_n \\ j'_1 & j'_2 & j'_3 & \cdots & j'_n \end{bmatrix}$$

zwei verschiedene Darstellungen einer Permutation σ;

$$f_1 := \{i_1, i_2, \ldots, i_n\}, \quad f_2 := \{j_1, j_2, \ldots, j_n\},$$

und

$$f'_1 := \{i'_1, i'_2, \ldots, i'_n\}, \quad f'_2 := \{j'_1, j'_2, \ldots, j'_n\}.$$

Wir mußen zeigen, dass

$$(-1)^{z(f_1)+z(f_2)} = (-1)^{z(f'_1)+z(f'_2)}.$$

Aus der Gleichung

$$\begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix} = \begin{bmatrix} i'_1 & i'_2 & i'_3 & \cdots & i'_n \\ j'_1 & j'_2 & j'_3 & \cdots & j'_n \end{bmatrix}$$

folgt, dass man

$$\begin{bmatrix} i'_1 & i'_2 & i'_3 & \cdots & i'_n \\ j'_1 & j'_2 & j'_3 & \cdots & j'_n \end{bmatrix}$$

durch Vertauschungen der Spalten von

$$\begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$$

bekommen kann. Nun genügt es zu bemerken: Ist z.B. $\begin{bmatrix} i'_1 & i'_2 & i'_3 & \cdots & i'_n \\ j'_1 & j'_2 & j'_3 & \cdots & j'_n \end{bmatrix}$ durch eine Vertauschung zweier Spalten von $\begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$ entstanden, so folgt aus 6.20, dass

$$(-1)^{z(f'_1)} = -(-1)^{z(f_1)},\ (-1)^{z(f'_2)} = -(-1)^{z(f_2)}$$

und damit

$$(-1)^{z(f_1)+z(f_2)} = (-1)^{z(f'_1)+z(f'_2)}.$$

□

Satz 6.22 *Seien $\sigma, \sigma' \in S_n$ beliebige Permutationen. Dann gilt*

$$sgn(\sigma \cdot \sigma') = sgn(\sigma) \cdot sgn(\sigma').$$

Beweis. Wir schreiben die Permutationen σ und σ' in der folgenden Gestahlt auf:

$$\sigma' = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$$

und

$$\sigma' = \begin{bmatrix} j_1 & j_2 & j_3 & \cdots & j_n \\ k_1 & k_2 & k_3 & \cdots & k_n \end{bmatrix}.$$

Dann erhalten wir

$$\sigma \cdot \sigma' = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ k_1 & k_2 & k_3 & \cdots & k_n \end{bmatrix}.$$

Seien

$$f_1 := \{i_1, i_2, \ldots, i_n\},\ \ f_2 := \{j_1, j_2, \ldots, j_n\},\ \ f_3 := \{k_1, k_2, \ldots, k_n\}.$$

Dann gilt

$$sgn(\sigma\sigma') = (-1)^{z(f_1)+z(f_3)} = (-1)^{z(f_1)+z(f_2)}(-1)^{z(f_2)+z(f_3)} = sgn(\sigma)sgn(\sigma').$$

□

Korollar 6.23

$$sgn(\sigma) = sgn(\sigma^{-1}).$$

7 Determinantentheorie

7.1 Determinante

Definition 7.1 Es sei $A = (a_{ij})$ eine quadratische $n \times n$-Matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Man betrachet ein Produkt von n Einträgen von A, so dass ein und nur ein Eintrag von jeder Zeile und ein und nur ein Eintrag von jeder Spalte herrührt. Ein solches Produkt kann in der Form

$$a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}$$

geschrieben werden, wobei

$$\sigma = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ j_1 & j_2 & j_3 & \cdots & j_n \end{bmatrix}$$

eine Permutation aus S_n ist. Die Permutation $\sigma \in S_n$ hängt nicht von der Reihenfolge der Faktoren im Produkt $a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}$ ab. Die **Determinante der Matrix** A, die durch $det(A)$ oder $|A|$ bezeichnet wird, ist die folgende Summe, die durch Aufsummieren aller Permutationen in S_n entsteht

$$|A| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Diese Formel nennt man die **Leibniz-Formel für die Determinante**.

Beispiel 7.2

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

Beispiel 7.3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Proposition 7.4 *Die Determinante einer Matrix A und die Determinante ihrer transponierten Matrix ${}^tA$ sind gleich: $|A| = |{}^tA|$.*

Beweis.

$$|\,{}^tA| = \sum_{\sigma\in S_n} sgn(\sigma)\,{}^ta_{1\sigma(1)}\,{}^ta_{2\sigma(2)}\cdots\,{}^ta_{n\sigma(n)}.$$

Wegen ${}^ta_{ij} = a_{ji}$ erhalten wir

$$|\,{}^tA| = \sum_{\sigma\in S_n} sgn(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}.$$

Nun bemerken wir, daß] die Permutation

$$\begin{bmatrix}\sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n)\\ 1 & 2 & 3 & \cdots & n\end{bmatrix}$$

inverse zu σ ist. Deswegen gilt

$$a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)}a_{2\sigma^{-1}(2)}\cdots a_{n\sigma^{-1}(n)}.$$

Wegen 6.23 gilt $sgn(\sigma) = sgn(\sigma^{-1})$ und wir erhalten

$$|\,{}^tA| = \sum_{\sigma\in S_n} sgn(\sigma^{-1})a_{1\sigma^{-1}(1)}a_{2\sigma^{-1}(2)}\cdots a_{n\sigma^{-1}(n)} = |A|,$$

weil die Permutation σ^{-1} alle Elemente von S_n durchläuft, wenn σ selbe alle Elemente von S_n durchläuft. □

7.2 Eigenschaften der Determinante

Proposition 7.5 *Sei $A \in M(n\times n;\mathbb{R})$ und A' die Matrix, die man aus A erhählt, wenn man zwei Zeilen $Z_i(A)$ und $Z_j(A)$ von A vertauscht. Dann gilt $|A'| = -|A|$.*

Beweis. Wir haben

$$|A'| = \sum_{\sigma\in S_n} sgn(\sigma)a'_{1\sigma(1)}a'_{2\sigma(2)}\cdots a'_{i\sigma(i)}\cdots a'_{j\sigma(j)}\cdots a'_{n\sigma(n)} =$$

$$\sum_{\sigma\in S_n} sgn(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{j\sigma(i)}\cdots a_{i\sigma(j)}\cdots a_{n\sigma(n)}.$$

Wegen 6.20 haben die Pernutationen

$$\sigma := \begin{bmatrix}1 & 2 & \cdots & i & \cdots & j & \cdots n\\ \sigma(1) & \sigma(2) & \cdots & \sigma(i) & \cdots & \sigma(j) & \cdots\sigma(n)\end{bmatrix}$$

und

$$\sigma' := \begin{bmatrix}1 & 2 & \cdots & j & \cdots & i & \cdots n\\ \sigma(1) & \sigma(2) & \cdots & \sigma(i) & \cdots & \sigma(j) & \cdots\sigma(n)\end{bmatrix}$$

verischiedene Vorzeichen: $sgn\,\sigma' = -sgn\,\sigma$. Deswegen gilt

$$|A'| = \sum_{\sigma' \in S_n} -sgn(\sigma') a_{1\sigma'(1)} a_{2\sigma'(2)} \cdots a_{i\sigma'(i)} \cdots a_{j\sigma'(j)} \cdots a_{n\sigma'(n)} = -|A|.$$

□

Korollar 7.6 *Sei $A \in M(n \times n; \mathbb{R})$. Besitzt A zwei indentisch gleiche Zeilen, so ist $|A| = 0$.*

Beweis. Bei Vertauschung der identischen Zeilen erhalten wir eine Matrix A', die gleich A ist. Andererseit wegen 7.5 gilt $|A'| = -|A|$. Aus $|A| = -|A|$ folgt $|A| = 0$. □

Proposition 7.7 *Sei $A \in M(n \times n; \mathbb{R})$ und A' die Matrix, die man aus A erhählt, wenn man die i-te Zeile $Z_i(A)$ von A mit einem Skalar λ multipliziert. Dann gilt $|A'| = \lambda |A|$.*

Beweis.

$$|A'| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots (\lambda a_{i\sigma(i)}) \cdots a_{n\sigma(n)} =$$

$$\lambda \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} = \lambda |A|.$$

□

Korollar 7.8 *Besteht eine Zeile von A aus Nullen, so ist $|A| = 0$.*

Proposition 7.9 *Sei $A \in M(n \times n; \mathbb{R})$. Ist die i-te Zeile von A als Summe zweier Vektoren dargestellt:*

$$Z_i(A) = (a_{i1}, \ldots, a_{in}) = (a'_{i1}, \ldots, a'_{in}) + (a''_{i1}, \ldots, a''_{in}),$$

so ist

$$|A| = |A'| + |A''|,$$

wobei

$$Z_j(A) = Z_j(A') = Z_j(A'') \;\; \forall j \neq i$$

und

$$Z_i(A') = (a'_{i1}, \ldots, a'_{in}), \;\; Z_i(A'') = (a''_{i1}, \ldots, a''_{in}).$$

Beweis.

$$|A| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} =$$

$$\sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \cdots a_{n\sigma(n)} =$$

$$\sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a'_{i\sigma(i)} \cdots a_{n\sigma(n)} +$$

$$+ \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a''_{i\sigma(i)} \cdots a_{n\sigma(n)} = |A'| + |A''|.$$

□

Proposition 7.10 *Sei $A \in M(n \times n; \mathbb{R})$ und A' die Matrix, die man aus A erhählt, wenn man λ-Vielefaches der j-ten Zeile von A zur i-ten Zeile addiert. Dann gilt $|A'| = |A|$.*

Beweis.

$$|A'| = \sum_{\sigma \in S_n} sgn(\sigma) a'_{1\sigma(1)} a'_{2\sigma(2)} \cdots a'_{i\sigma(i)} \cdots a'_{j\sigma(j)} \cdots a'_{n\sigma(n)} =$$

$$\sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots (a_{i\sigma(i)} + \lambda a_{j\sigma(j)}) \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} =$$

Wegen 7.9 und 7.7 erhalten wir:

$$|A'| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} +$$

$+$

$$\lambda \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{j\sigma(j)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)}.$$

Der zweite Summand ist wegen 7.6 zu Null gleich. Der erste Summand ist zu $|A|$ gleich. □

Proposition 7.11 *Sei $A \in M(n \times n; \mathbb{R})$. Ist $a_{12} = a_{13} = \ldots = a_{1n} = 0$, so ist*

$$|A| = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} & \ldots & a_{2n} \\ a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n2} & a_{n3} & \ldots & a_{nn} \end{vmatrix}$$

Bewies. Die Matrix A sieht wie folg aus:

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

Wegen $a_{1k} = 0 \ \forall k > 1$ ist ein Produkt $a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)}$ in der Leibniz-Formel

$$|A| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)}$$

nur dann ungleich Null sein kann, wenn $\sigma(1) = 1$ ist. Die Teilmenge

$$\{\sigma \in S_n \ : \ \sigma(1) = 1\}$$

besitzt $(n-1)!$ Elemente und kann mit der Menge der Permutationen $S(X)$, $X := \{2, \dots, n\}$ identifiziert werden. Diese Identifikation ergibt sich durch die folgende Abbildung:

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ 1 & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{bmatrix} \mapsto \sigma' := \begin{bmatrix} 2 & 3 & \cdots & n \\ \sigma(2 & \sigma(3) & \cdots & \sigma(n) \end{bmatrix} \in S(X).$$

Ausserdem gilt $sgn(\sigma) = sgn(\sigma')$. Daher folgt

$$|A| = \sum_{\sigma \in S_n, \sigma(1)=1} sgn(\sigma) a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} =$$

$$= a_{11} \sum_{\sigma' \in S(X)} sgn(\sigma') \prod_{x \in X} a_{x\sigma'(x)} = a_{11}|A'|,$$

wobei

$$A' = \begin{pmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

ist.

□

Proposition 7.12 *Ist $A \in M(n \times n; \mathbb{R})$ eine untere Dreicksmatrix, d.h. die oberhalb der Diagonale nur Nullen hat, so ist $|A|$ gleich dem Produkt der Diagonalelemente. Insbesondere gilt $|E_n| = 1$.*

Beweis. Folgt aus 7.11 mit Induktion nach n. □
Aus 7.4 und 7.12 folgt sofort:

Korollar 7.13 *Ist $A \in M(n \times n; \mathbb{R})$ eine obere Dreicksmatrix, d.h. die unterhalb der Diagonale nur Nullen hat, so ist $|A|$ gleich dem Produkt der Diagonalelemente.*

Definition 7.14 Sei $A \in M(n \times n; \mathbb{R})$. Wir bezeichnen mit A_{ij} die $(n-1) \times (n-1)$-Matrix, die man durch Streichen der i-ten Zeile und der j-ten Spalte aus A erhält.

Satz 7.15 (Entwicklungssatz von Laplace) *Sei $A \in M(n \times n; \mathbb{R})$. Für jedes $i \in \{1, \ldots, n\}$ gilt*

$$|A| = a_{i1}(-1)^{i+1}|A_{i1}| + a_{i2}(-1)^{i+2}|A_{i2}| + \cdots + a_{in}(-1)^{i+n}|A_{in}| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}|.$$

Diese Formel heißt **Entwicklung von $|A|$ nach i-ten Zeile**.

Beweis. Wir schreiben $Z_i(A)$ als Linearkombination auf:

$$Z_i(A) = (a_{i1}, \ldots, a_{in}) = a_{i1}(1, 0, \ldots, 0) + \cdots + a_{in}(0, 0, \ldots, 1).$$

Wegen 7.9 und 7.7 gilt

$$|A| = a_{i1}|B_1| + \cdots + a_{in}|B_n| = \sum_{j=1}^{n} a_{ij}|B_j|,$$

wobei

$$Z_k(B_j) = Z(A) \;\; \forall j, \forall k \neq i$$

und

$$Z_i(B_j) = (0, \ldots, \underbrace{1}_{j}, \ldots, 0).$$

Nach $i-1$ Zeilenvertauschungen und $j-1$ Spaltenvertauschnungen bekommt man aus B_j die $n \times n$-Matrix

$$\begin{pmatrix} 1 & 0 \\ * & A_{ij} \end{pmatrix}.$$

Wegen 7.11 und 7.5 erhalten wir

$$|B_j| = (-1)^{i-1}(-1)^{j-1}|A_{ij}| = (-1)^{i+j}|A_{ij}|.$$

Zusammen mit $\sum_{j=1}^{n} a_{ij}|B_j|$ dies liefert uns die gewünschte Formel

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}|.$$

□

Aus 7.4 und 7.15 folgt sofort:

Korollar 7.16 *Sei $A \in M(n \times n; \mathbb{R})$. Für jedes $j \in \{1, \ldots, n\}$ gilt*

$$|A| = a_{1j}(-1)^{1+j}|A_{1j}| + a_{2j}(-1)^{2+j}|A_{2j}| + \cdots + a_{nj}(-1)^{n+j}|A_{nj}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}|.$$

Diese Formel heißt **Entwicklung von $|A|$ nach j-ten Spalte**.

7.3 Anwendungen der Determinante

Satz 7.17 *Sei $A \in M(n \times n; \mathbb{R})$. Dann gilt: $Rang(A) = n \;\Leftrightarrow |A| \neq 0$.*

Beweis. Sei A' die Matrix in der Zeilenstufenform, die aus A durch elementare Zeilenumformungen vom 1. und 2. Typ entstanden ist. Wegen 5.19 erhalten wir $Rang(A) = Rang(A')$. Andererseits folgt aus 7.5 und 7.10, dass $|A'| = \pm|A|$ ist. Deswegen genügt es die Aussage nur für A' zu beweisen.
Die Matrix A' hat Rang n genau dann, wenn A' genau n von Null verschiedenen Zeilen hat. Das letzte gilt genau dann, wenn $a'_{ii} \neq 0 \quad \forall i \in \{1, \ldots, n\}$, d.h. $a'_{11}a'_{22}\cdots a'_{nn} = |A'| \neq 0$ (s. 7.13).

□

Satz 7.18 *Sei $A \in M(n \times n; \mathbb{R})$. Die folgenden Bedingungen sind äquivalent:*
(1) $|A| \neq 0$;
(2) $Rang(A) = n$;
(3) die lineare Abbildung $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ ist surjektiv;
(4) die lineare Abbildung $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ ist injektiv;
(5) die lineare Abbildung $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ ist bijektiv;
(6) die inverse Matrix A^{-1} existiert.

Beweis. (1) $\Leftrightarrow$ (2) ist schon in 7.17 bewiesen.
(2) $\Leftrightarrow$ (3) und (2) $\Leftrightarrow$ (4) folgen aus 6.14. Damit ist die Bedingung (2) zu (5) äquivalent.
Aus 6.6 und 6.11 folgt, dass (6) zu (5) äquivalent. □

Definition 7.19 Sei $A \in M(n \times n; \mathbb{R})$. Mit $A^{\#}$ bezeichnen wir die Matrix mit Einträgen

$$a^{\#}_{ij} = (-1)^{i+j}|A_{ji}|.$$

Die Matrix $A^{\#}$ wird zu A **komplementäre** (oder **adjungierte**) Matrix genannt.

Satz 7.20 (Formel für inverse Matrix) *Sei $A \in M(n \times n; \mathbb{R})$. Dann gilt*

$$A \cdot A^{\#} = A^{\#} \cdot A = |A| \cdot E_n.$$

Insbesondere, hat man

$$A^{-1} = \frac{1}{|A|} A^{\#},$$

wenn $|A| \neq 0$.

Beweis. Sei $C = A \cdot A^{\#}$. Dann gilt

$$c_{ij} = \sum_{k=1}^{n} a_{ik} a_{kj}^{\#} = \sum_{k=1}^{n} a_{ik} (-1)^{k+j} |A_{jk}|.$$

Ist $i = j$, so ist wegen 7.15

$$c_{ii} = \sum_{k=1}^{n} a_{ik} (-1)^{k+i} |A_{ik}| = |A|.$$

Ist $i \neq j$, so kann man wegen 7.15 die Summe

$$\sum_{k=1}^{n} a_{ik} (-1)^{k+j} |A_{jk}|$$

als Entwicklung nach j-ten Zeile der Determinante einer Matrix interpretieren, derer i-te und j-te Zeilen identisch sind. Wegen 7.6 ist diese Determinante gleich Null, d.h. $c_{ij} = 0 \ \forall i \neq j$.

□

Satz 7.21 (Die Cramersche Regel) *Sei* $A \in M(n \times n; \mathbb{R})$ *und* $b \in \mathbb{R}^n$ *ein beliebiger Spaltenvektor. Dann gilt:*
(1) Das lineare Gleichungssystem $Ax = b$ *besitzt eine einzige Lösung genau dann, wenn* $|A| \neq 0$.
(2) Ist $|A| \neq 0$, *so ist die Lösung von LGS* $Ax = b$ *durch die folgenden Formeln bestimmt*

$$x_i = \frac{|B_i|}{|A|}, \tag{3}$$

wobei die Matrix $B_i \in M(n \times n; \mathbb{R})$ *durch Ersetzung mit dem Vektor* b *der* i*-ten Spalte von* A *entsteht, d.h.*

$$S_j(B_i) = S_j(A) \ \ \forall j \neq i,$$

$$S_i(B_i) = b.$$

Die Formel (3) heißt die **Cramersche Regel**.

Beweis. Die erste Aussage folgt sofort aus der Äquivalenz (1) und (5) in 7.18. Für die zweite Aussage schreiben wir die Lösung x als $A^{-1}b$ auf:

$$x_i = \langle Z_i(A^{-1}), b\rangle, \quad 1 \leq i \leq n.$$

Wegen 7.20, haben wir

$$x_i = \sum_{j=1}^{n}(-1)^{i+j}\frac{|A_{ji}|}{|A|}b_j = \frac{1}{|A|}\sum_{j=1}^{n}(-1)^{i+j}|A_{ji}|b_j.$$

Andererseits kann die Summe

$$\sum_{j=1}^{n}(-1)^{i+j}|A_{ji}|b_j$$

als Entwicklung der Determinante von B_i nach i-ten Spalte enterpretiert werden. Damit erhalten wir

$$x_i = \frac{|B_i|}{|A|}.$$

□

7.4 Weitere Eigenschaften der Determinante

Satz 7.22 *Sei $D \in M((m+n)\times(m+n), \mathbb{R})$ eine Matrix der Form*

$$D = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix},$$

wobei $A \in M(m\times m, \mathbb{R})$, $B \in M(n\times n, \mathbb{R})$, und $C \in M(m\times n, \mathbb{R})$. Dann gilt

$$|D| = |A|\cdot|B|.$$

Beweis. Durch elementare Zeilenumformungen vom Typ 1 und 2 kann man die Matrizen $(A\ C)$ und $(0\ B)$ in die Zeilenstufenformen $(A'\ C')$ und $(0\ B')$ über führen. Wir bezeichnen mit D' die Matrix

$$D' := \begin{pmatrix} A' & C' \\ 0 & B' \end{pmatrix},$$

die aus D durch diese Zeilneumformungen entsteht.
Sei k (bzw. l) die Anzahl Zeilenumformungen vom Typ 1, die man bei der Überführung $(A\ C) \to (A'\ C')$ (bzw. $(0\ B) \to (0\ B')$) braucht. Wegen 7.5 erhalten wir

$$|A'| = (-1)^k|A|, \quad |B'| = (-1)^l|B|, \quad |D'| = (-1)^{k+l}|D|.$$

Andererseits sind A', B', D' die oberen Dreicksmatrizen. Deswegen gilt

$$|A'| = a'_{11} \cdots a'_{mm},\ |B'| = b'_{11} \cdots b'_{nn},\ |D'| = d'_{11} \cdots d'_{(m+n)(m+n)}.$$

Wir bemerken nun, dass

$$d'_{11} = a'_{11}, \ldots, d'_{mm} = a'_{mm},\ \ d'_{(m+1)(m+1)} = b'_{11}, \ldots, d'_{(m+n)(m+n)} = b'_{nn}.$$

Deswegen gilt $|D'| = |A'||B'|$ und damit

$$(-1)^{k+l}|D| = (-1)^k|A| \cdot (-1)^l|B|.$$

Also gilt $|D| = |A| \cdot |B|$.

□

Korollar 7.23 *Sei $D \in M((m+n) \times (m+n), \mathbb{R})$ eine Matrix der Form*

$$D = \begin{pmatrix} A & 0 \\ C & B \end{pmatrix},$$

wobei $A \in M(m \times m, \mathbb{R})$, $B \in M(n \times n, \mathbb{R})$, und $C \in M(n \times m, \mathbb{R})$. Dann gilt

$$|D| = |A| \cdot |B|.$$

Beweis. Es gilt

$${}^tD = \begin{pmatrix} {}^tA & {}^tC \\ 0 & {}^tB \end{pmatrix}.$$

Aus 7.22 folgt, dass $|{}^tD| = |{}^tA||{}^tB|$. Wegen 7.4 erhalten wir $|{}^tD| = |D|$, $|{}^tA| = |A|$, $|{}^tB| = |B|$ und damit $|D| = |A| \cdot |B|$. □

Satz 7.24 *Seien $A, B \in M(n \times n; \mathbb{R})$ beliebige quadratische Matrizen. Dann gilt*

$$|AB| = |A||B|.$$

Beweis. Sei

$$D := \begin{pmatrix} A & 0 \\ -E_n & B \end{pmatrix}.$$

Wegen 7.22 und 7.4 erhalten wir

$$|D| = |A| \cdot |B|.$$

Nun werden wir D mit elementaren Zeilenumformungen von Typ 2 ändern:

Für alle Paare (i,j) $(1 \leq i,j \leq n)$ wir addieren die a_{ij}-fache $(n+j)$-te Zeile von D zur i-ten Zeile von D. Durch diese n^2 elementare Zeilenumformungen vom Typ 2 erreichen wir die Matrix

$$D' := \begin{pmatrix} 0 & C \\ -E_n & B \end{pmatrix},$$

wobei $C = AB$ ist. Wegen 7.10 gilt $|D'| = |D| = |A||B|$. Andererseits kann man durch n Vertauschungen der Spalten von D die Matrix

$$D'' := \begin{pmatrix} AB & 0 \\ B & -E_n \end{pmatrix}$$

bekommen. Wegen 7.22 und 7.4 gilt $|D''| = |AB| \cdot |-E_n| = (-1)^n |AB|$. Wegen 7.5 gilt $|D''| = (-1)^n |D'| = (-1)|D|$. Damit erhalten wir

$$|AB| = |A| \cdot |B|.$$

□

7.5 Vandermondesche Determinate

Definition 7.25 Seien $x_1, \ldots, x_n$ reelle Zahlen. Mit $\Delta_n = \Delta_n(x_1, \ldots, x_n)$ bezeichnen wir die Determinate der $n \times n$-Matrix:

$$\Delta_n := \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}$$

Δ_n heißt **Vandermondesche Determinante**.

Satz 7.26

$$\Delta_n = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Beweis. Induktion über n. Induktionsanfang: Für $n = 2$ erhalten wir

$$\Delta_2 = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = (x_2 - x_1).$$

Induktionsschluß: Die Formel gelte für $n-1$. Wir ziehen nun von jeder Spalte (und zwar angefangen von der letzten bis zur zweiten) jeweils die mit x_1 multiplizierte

vorausgehende Spalte ab. Bis auf das erste Element werden dadurch alle Elemente der ersten Zeile zu 0 und man erhält

$$\Delta_n = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1 x_2 & \cdots & x_2^{n-1} - x_1 x_2^{n-2} \\ 1 & x_3 - x_1 & x_3^2 - x_1 x_3 & \cdots & x_3^{n-1} - x_1 x_3^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n - x_1 & x_n^2 - x_1 x_n & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix} =$$

$$= \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1 x_2 & \cdots & x_2^{n-1} - x_1 x_2^{n-2} \\ x_3 - x_1 & x_3^2 - x_1 x_3 & \cdots & x_3^{n-1} - x_1 x_3^{n-2} \\ \cdots & \cdots & \cdots & \cdots \\ x_n - x_1 & x_n^2 - x_1 x_n & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix}.$$

Jetzt kann man aus der i-ten Zeile ($i = 1, \ldots, n-1$) den Faktor $x_{i+1} - x_1$ herausziehen. Das ergibt:

$$\Delta_n = \prod_{i=2}^{n} (x_i - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} \end{vmatrix} = \left(\prod_{i=2}^{n} (x_i - x_1) \right) \Delta_{n-1}(x_2, x_3, \ldots, x_n).$$

Nach Induktionsvoraussetzung gilt

$$\Delta_{n_1}(x_2, x_3, \ldots, x_n) = \prod_{2 \le k < j \le n} (x_j - x_k).$$

Damit erhalten wir

$$\Delta_n = \prod_{i=2}^{n} (x_i - x_1) \cdot \prod_{2 \le k < j \le n} (x_j - x_k) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

□

Korollar 7.27 *Sind $x_1, \ldots, x_n$ paarweise verschieden, so ist $\Delta_n(x_1, \ldots, x_n) \neq 0$.*

Seien $x_1, \ldots, x_n$ paarweise verschiedene reelle Zahlen. Wir werden eine polynomiale Funktion

$$f(t) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1}$$

von Grad $\le n - 1$ suchen, die gegebene Werte $y_1, \ldots, y_n$ in den Punkten $x_1, \ldots, x_n$ hat: $y_i = f(x_i)$ $1 \le i \le n$.

Satz 7.28 *Sind $x_1, \ldots, x_n$ paarweise verschiedene reelle Zahlen und $y_1, \ldots, y_n$ beliebige reelle Zahlen, so gibt es genau ein Polynom $f(t)$ vom Grad $\leq n-1$, so dass gilt $y_i = f(x_i)$ $1 \leq i \leq n$. Das Polynom $f(t)$ läßt sich durch die folgenede explizite Formel schreiben:*

$$f(t) = \sum_{i=1}^{n} y_i \frac{\prod_{j \neq i}(t - x_j)}{\prod_{j \neq i}(x_i - x_j)}.$$

Diese Formel heißt **Lagrangesche Interpolationformel**.

Beweis. Die Gleichungen $y_i = f(x_i)$ $(1 \leq i \leq n)$ kann man in Matrizenform wie folgt schreiben:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ y_n \end{pmatrix}$$

Wegen 7.27 besitzt dieses lineare Gleichungssystem eine eindeutige Lösung. Also exisitiert ein einziges Polynom $f(t)$ vom Grad $\leq n-1$, so dass gilt $y_i = f(x_i)$ $(1 \leq i \leq n)$.
Wir bezeichnen nun mit $g_i(t)$ $(1 \leq i \leq n)$ das Polynom vom Grad $n-1$

$$g_i(t) := \frac{\prod_{j \neq i}(t - x_j)}{\prod_{j \neq i}(x_i - x_j)}.$$

Offensichtlich gilt $g_i(x_i) = 1$ und $g_i(x_j) = 0$, falls $j \neq i$. Das Polynom $f(t) := y_1 g_1 + \cdots y_n g_n$ hat Grad $\leq n-1$ und es gilt $f(x_i) = y_i g(x_i) = y_i$ $(1 \leq i \leq n)$. Deswegen ist $f(t)$ genau das einzige gesuchte Polynom. □

7.6 Determinante als multilineare alternierende Funktion

Definition 7.29 Sei $F : M(n \times n, \mathbb{R}) \to \mathbb{R}$ eine Funktion, die jeder Matrix $A \in M(n \times n, \mathbb{R})$ eine reelle Zahl $F(A)$ zuordnet. Wir werden diese Funktion als eine Funktion von Zeilen $Z_1(A), \ldots, Z_n(A)$ auffassen und schreiben sie in der Form $F(Z_1(A), \ldots, Z_n(A))$ auf. Die Function F heißt **multilinear**, wenn
(1) für jedes $i \in \{1, \ldots, n\}$ gilt:

$$F(Z_1(A), \ldots, Z_i'(A) + Z_i''(A), \ldots, Z_n(A)) =$$

$$= F(Z_1(A), \ldots, Z_i'(A), \ldots, Z_n(A)) + F(Z_1(A), \ldots, Z_i''(A), \ldots, Z_n(A));$$

(2) für jedes $i \in \{1, \ldots, n\}$ gilt:

$$F(Z_1(A), \ldots, \lambda Z_i(A), \ldots, Z_n(A)) = \lambda F(Z_1(A), \ldots, Z_i(A), \ldots, Z_n(A)),\ \lambda \in \mathbb{R}.$$

Die Funktion F heißt **alternierend**, wenn für alle $1 \leq i < j \leq n$ gilt:

$$F(Z_1(A), \ldots, Z_i(A), \ldots, Z_j(A), \ldots, Z_n(A)) =$$

$$-F(Z_1(A), \ldots, Z_j(A), \ldots, Z_i(A), \ldots, Z_n(A)).$$

Satz 7.30 *Jede multilineare alternierende Funktion* $F : M(n \times n, \mathbb{R}) \to \mathbb{R}$ *hat Gestalt*

$$F(A) = F(E_n) \cdot |A|.$$

Isbesondere ist Determinante als multilineare alternierende Funktion F *eindeutig durch die Bedingung* $F(E_n) = 1$ *bestimmt.*

Beweis. Durch elementare Zeilenumformungen kann man A in einer Matrix in normierten Zeilenstufenform überführen. Bei jeder Zeilenumformung verhält sich die Funktion F genau wie Determinante. Andererseits: Ist A schon in der normierten Zeilenstufenform, so ist die Aussage des Satzes offensichtlich. □

Korollar 7.31

$$|AB| = |A||B|.$$

Beweis. Man betrachten die Funktion

$$F : A \mapsto |AB|.$$

Wir zeigen, dass diese Funktion multilinear und alternierend ist. Es sei $Z_i(A)$ als Summe $Z_i(A) = Z_i'(A) + Z_i''(A)$ dargestellt. Dann gilt

$$Z_i(AB) = Z_i(A)B = (Z_i'(A) + Z_i''(A))B = Z_i'(A)B + Z_i''(A)B.$$

Wir bezeichnen mit A' und A'' die quadratischen $n \times n$-Matrizen, wobei $Z_j(A) = Z_j(A') = Z_j(A'')$ für $j \neq i$ und $Z_i(A') = Z_i'(A)$ und $Z_i(A'') = Z_i''(A)$ sind. Dann erhalten wir

$$Z_i(AB) = Z_i(A'B) + Z_i(A''B).$$

Deswegen gilt

$$F(A) = |AB| = |A'B| + |A''B| = F(A') + F(A'').$$

Ist $Z_i(A) = \lambda Z_i'(A)$, so ist

$$Z_i(AB) = Z_i(A)B = (\lambda Z_i'(A))B = \lambda(Z_i'(A)B = \lambda Z_i(A'B),$$

wobei $Z_j(A) = Z_j(A')$ für $j \neq i$ und $Z_i(A') = Z_i(A)'$ ist. Deswegen gilt

$$F(A) = |AB| = \lambda |A'B| = \lambda F(A').$$

Also ist F multilinear.
Wegen 4.11 entsteht die Matrix $I_{ij}A$ durch Vertauschung der i-ten und j-ten Zeile in der Matrix A. Andererseits entsteht die Matrix $I_{ij}AB$ durch Vertauschung der i-ten und j-ten Zeile in der Matrix AB. Damit erhalten wir

$$F(I_{ij}A) = |I_{ij}AB| = -|AB| = -F(A).$$

Also ist F alternierend.
Wegen 7.30 gilt

$$F(A) = F(E_n) \cdot |A|.$$

Andererseits haben wir $F(E_n) = |E_nB| = |B|$. Deswegen gilt

$$|AB| = F(A) = |B||A| = |A||B|.$$

□

7.7 Minoren und Rang

Definition 7.32 Sei $A \in M(m \times n, \mathbb{R})$. Ist $\{i_1, \ldots, i_k\}$ eine Teilmenge von $\{1, \ldots, m\}$ und $\{j_1, \ldots, j_k\}$ eine Teilmenge von $\{1, \ldots, n\}$, so bezeichnen wir mit $A^{j_1,\ldots,k}_{i_1,\ldots,i_k}$ die Untermatrix in A, deren Einträgen in den Zeilen $Z_{i_1}(A), \ldots, Z_{i_k}(A)$ und in den Spalten $S_{j_1}(A), \ldots, S_{j_k}(A)$. Die Determinante $|A^{j_1,\ldots,j_k}_{i_1,\ldots,i_k}|$ der quadratischen $k \times k$-Matrix $A^{j_1,\ldots,j_k}_{i_1,\ldots,i_k}$ wird **Minor k-ter Ordnung** von A genannt.

Satz 7.33 *Die Zahl r ist zum Rang von $A \in M(m \times n, \mathbb{R})$ gleich genau dann, wenn einen von Null verschiedenen Minor der r-ten Ordnung gibt, und alle Minoren der $(r+1)$-ten Ordnung gleich Null sind.*

Beweis. Sei $A' \in M(m \times n, \mathbb{R})$ aus A durch endlich viele Zeilen- und Spaltenumformungen entstandene Matrix. Dann sind alle Minoren der k-ten Ordnung von A' als Linearkombinationen von Minoren der k-ten Ordnung von A darstellbar. Andererseits kann man durch Zeilen- und Spaltenumformungen immer die Matrix $A'' = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ erreichen, wobei $r = Rang(A)$ ist. Für die Matrix A'' ist die Aussage offensichtlich. Da die Matrix A aus A'' durch elementare Umformungen entsteht, gilt die Aussage auch für A.

□

8 Gruppen, Ringe und Körpern

8.1 Gruppen

Definition 8.1 Eine **Gruppe** ist ein Paar $(G, \circ)$ bestehend aus einer Menge G und einer Verknüpfung $\circ$:

$$G \times G \to G$$

$$(x, y) \mapsto x \circ y,$$

so dass die folgenden Axiome erfüllt sind:
(i) **Assoziativgesetz**:

$$x \circ (y \circ z) = (x \circ y) \circ z \quad \forall x, y, z \in G.$$

(ii) Es gibt ein Element $e \in G$, so dass $\forall x \in G$ gilt:

$$x \circ e = e \circ x = x.$$

(Ein Element e ist eindeutig bestimmt, denn ist e' ein weiteres solches Element, so gilt $e = e \circ e' = e'$. Man nennt e das **neutrale Element** von G.)
(iii) Zu jedem $x \in G$ gibt es ein $y \in G$, so dass gilt:

$$x \circ y = y \circ x = e.$$

(y ist ebenfalls eindeutig bestimmt, denn ist $x \circ y' = y' \circ x = e$ für ein anderes $y' \in G$, so folgt

$$y = y \circ e = y \circ (x \circ y') = (y \circ x) \circ y' = e \circ y' = y'.$$

Man nennt y das **Inverse** von x. Es wird mit x^{-1} bezeichnet.

Bezeichnung 8.2 (i) Es ist üblich, das Verknüpfungzeichen $\circ$ wegzulassen und kurz xy statt $x \circ y$ schreiben.
(ii) Wegen des Assoziativgesetzes darf man Klammern in mehrfachen Produkten weglassen, z. B. ist

$$xyz = (xy)z = x(yz)$$

$$xyzw = (xyz)w = (xy)(zw) = \cdots$$

Man darf aber im allgemeinen nicht die Reihenfolge der Faktoren ändern.

Definition 8.3 Ist $xy = yx \ \forall x, y \in G$, so heißt die Gruppe G **kommutativ** oder eine **abelsche Gruppe**.

Bezeichnung 8.4 Die Verknüpfung in abelschen Gruppen wird oftmals additiv geschrieben. Statt $x \circ y$ (oder xy) schreibt man also $x + y$, statt x^{-1} schreibt man $-x$, und das neutrale Element e bezeichnet man mit 0.

Beispiel 8.5 Die Menge der Matrizen $A \in M(n \times n, \mathbb{R})$ mit $|A| \neq 0$ bezeichnen wir mit $GL(n, \mathbb{R})$. Offensichtlich ist $GL(n, \mathbb{R})$ eine Gruppe bezüglich der Multiplikation. Diese Gruppe heißt **allgemeine lineare Gruppe**. Ist $n = 1$, so ist $GL(n, \mathbb{R}) = \mathbb{R}^* := \mathbb{R} \setminus \{0\}$.

Beispiel 8.6 Die Menge der Permutationen S_n bildet auch eine Gruppe. Diese Gruppe heißt **symmetrische Gruppe**.

Definition 8.7 Ist $(G, \circ)$ eine Gruppe, so sind die **Potenzen x^n eines Elementes** $x \in G$ $\forall n \in \mathbb{Z}$ folgendermaßen definiert:

$$x^0 = e$$

$$x^n = x(x^{n-1}) \text{ für } n \geq 1$$
$$x^{-n} = (x^{-1})^n \text{ für } n \geq 1.$$

Man zeigt dann leicht die Rechenregeln:

Proposition 8.8 *Für jedes $x \in G$ gilt*

$$x^{n+m} = x^n x^m,\ (x^m)^n = x^{nm} = (x^n)^m \ \ \forall n, m \in \mathbb{Z}$$

und

$$(xy)^{-1} = y^{-1}x^{-1} \ \ \forall x, y \in G.$$

Ist $xy = yx$, so ist

$$(xy)^n = x^n y^n \ \ \forall n \in \mathbb{Z}$$

Definition 8.9 Sei g ein Element aus G. Die minimale Zahl $k \in \mathbb{N}$, so dass $g^k = e$ ist, heißt die **Ordnung von** g und wird mit $Ord\, g$ bezeichnet. Man sagt, dass $Ord\, g = \infty$, wenn $g^k \neq e$ für alle $k \in \mathbb{N}$.

Definition 8.10 Eine Gruppe G heißt **zyklisch**, wenn es ein Element $g \in G$ gibt, so dass

$$G = \{g^n : n \in \mathbb{Z}\}.$$

In dieser Situation wird G mit $< g >$ bezeichnet. Das Element g nennt man einen **Erzeuger** von $G =< g >$.

Definition 8.11 Sei m eine positive ganze Zahl. Wir teilen $\mathbb{Z}$ in m **Klassen** folgendermaßen ein: Zu jedem $r \in \{0, 1, \ldots, m-1\}$ betrachten wir die Menge

$$r + m\mathbb{Z} := \{r + m \cdot n \ : \ n \in \mathbb{Z}\}.$$

Man hat die folgende disjunkte Vereinigung:

$$\mathbb{Z} = (0 + m\mathbb{Z}) \cup (1 + m\mathbb{Z}) \cup \cdots \cup (m - 1 + m\mathbb{Z}).$$

Diese Klassen nennt man **Restklassen modulo** m. Jede Klasse besteht aus allen Zahlen, die bei Division dutch m den gleichen Rest hinterlassen. Zwei Zahlen k und k' liegen genau dann in derselben Restklasse, wenn $k - k'$ durch m teilbar ist. Zu jedem $a \in \mathbb{Z}$ bezeichnen wir mit $\overline{a} = a + m\mathbb{Z}$ die Restklasse von a. Wir definieren die Addition

$$\overline{a} + \overline{b} := \overline{a + b}$$

und bemerken, dass diese Definition nicht von der Auswahl der Representanten abhängt.

Proposition 8.12 *Ist $m \in \mathbb{Z}$ und $m > 0$, so ist die Menge*

$$\mathbb{Z}/m\mathbb{Z} := \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$$

der Restklassen modulo m mit der oben erklärten Addition eine abelsche Gruppe.

Definition 8.13 Es sei $(G, \circ)$ eine Gruppe. Eine nichtleere Teilmenge $H \subset G$ heißt **Untergruppe** von G, wenn gilt:
(i) $x, y \in H \Rightarrow xy \in H$
(ii) $x \in H \Rightarrow x^{-1} \in H$.

Bemerkung 8.14 Ist $H \subset G$ eine Untergruppe von G, so gilt $e \in H$ (wegen (i), (ii) und $e = x \circ x^{-1}$). Eine Untergruppe H ist mit der von G induzierten Verknüpfung (oder Multiplikation) natürlich ebenfalls eine Gruppe.

Definition 8.15 Sei g ein Element aus G. Mit $< g >$ bezeichnen wir auch die **zyklische Untergruppe** mit dem Erzeuger g :

$$< g >= \{g^n \ : \ n \in \mathbf{Z}\} \subset G.$$

Beispiel 8.16 Die Teilmenge

$$SL(n, \mathbb{R}) := \{A \in GL(n, \mathbb{R}) \ : \ |A| = 1\} \subset GL(n, \mathbb{R})$$

ist eine Untergruppe, die **spezielle lineare Gruppe** heißt.

Definition 8.17 Es seien $(G, \circ)$ und $(G', \circ')$ Gruppen. Eine Abbildung

$$\varphi \ : \ G \to G'$$

heißt **Homomorphismus** (oder **Gruppenhomomorphismus**), wenn gilt:

$$\varphi(x \circ y) = \varphi(x) \circ' \varphi(y) \ \ \forall x, y \in G.$$

Proposition 8.18 *Sei $\varphi \ : \ G \to G'$ ein Gruppenhomomorphismus, e sei neutrale Element von G, e' das von G'. Dann gilt:*
(i) $\varphi(e) = e'$;
(ii) $\varphi(x^{-1}) = (\varphi(x))^{-1}$.

Beweis. (i) Nach Definitionen von φ und e, gilt:

$$\varphi(e) \circ' \varphi(e) = \varphi(e \circ e) = \varphi(e).$$

Wir multiplizieren die Gleichung $\varphi(e) \circ' \varphi(e) = \varphi(e)$ mit $(\varphi(e))^{-1}$ und erhalten damit $\varphi(e) = e'$.
(ii) Aus den Definitionen von φ, x^{-1} und aus der Aussage (i) folgt:

$$\varphi(x) \circ' \varphi(x^{-1}) = \varphi(x \circ x^{-1}) = \varphi(e) = e'.$$

Andererseits

$$\varphi(x^{-1}) \circ' \varphi(x) = \varphi(x^{-1} \circ x)\varphi(e) = e'.$$

Deshalb ist $\varphi(x^{-1})$ das Inverse von $\varphi(x)$. □

Definition 8.19 Einen Homomorphismus $\varphi : G \to G'$ heißt
(i) **Epimorphismus**, wenn φ surjektiv ist;
(ii) **Monomorphismus**, wenn φ injektiv ist;
(iii) **Isomorphismus**, wenn φ bijektiv ist.

Definition 8.20 Zwei Gruppen G und G' heißen **isomorph**, wenn es einen Gruppenisomorphismus $\varphi \ : \ G \to G'$ gibt. Man schreibt dann $G \cong G'$.

Definition 8.21 Sei $\varphi \ : \ G \to G'$ ein Gruppenhomomorphismus, e' das neutrale Element in G'. Die Untergruppe $Ker\,\varphi := \varphi^{-1}(e')$ heißt der **Kern** von φ.

Proposition 8.22 *Sei $\varphi \ : \ G \to G'$ ein Homomorphismus. Dann sind das Bild $Im\,\varphi := \varphi(G)$ und der Kern $Ker\,\varphi := \varphi^{-1}(e')$ Untergruppen von G'.*

Beweis. Seien x', y' beliebige Elemente aus $\varphi(G)$. Dann existieren Elemente $x, y \in G$, so dass $\varphi(x) = x'$ und $\varphi(y) = y'$. Daraus folgt, dass $x'y' = \varphi(x)\varphi(y) = \varphi(xy)$ in $\varphi(G)$ liegt.
Sei x' ein beliebiges Element aus $\varphi(G)$. Dann gibt es $x \in G$, so dass $\varphi(x) = x'$. Nach 8.18(ii), ist $\varphi(x^{-1})$ das Inverse von $\varphi(x)$. Deshalb gehört $(x')^{-1}$ auch zu $\varphi(G)$.

□

Beispiel 8.23 Die Exponentialabbildung

$$\exp \ : \ \mathbb{R} \to \mathbb{R}^*$$

ist ein Homomorphismus von $(\mathbb{R}, +)$ nach $(\mathbb{R}^*, \cdot)$:

$$\exp(x + y) = \exp(x) \cdot \exp(y)$$

Dieser Homomorphismus ist injektiv, aber nicht surjektiv (kein Epimorphismus).

Beispiel 8.24 Der Logarithmus

$$\log : \ \mathbb{R}_{>0} \to \mathbb{R}$$

ist ein Homomorphismus (sogar ein Isomorphismus) von $(\mathbb{R}_{>0}, \cdot)$ nach $(\mathbb{R}, +)$:

$$\log(xy) = \log(x) \log(y).$$

Beispiel 8.25 Die Determinante

$$\det \ : \ GL(n, \mathbb{R}) \to \mathbb{R}^*$$

ist ein Epimorphismus:

$$\det(AB) = \det A \cdot \det B.$$

Der Epimorphismus det is genau dann bijektiv, wenn $n = 1$.

Beispiel 8.26 Die Abbildung

$$\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, \ \ a \mapsto \overline{a}$$

ist ein Gruppenhomomorphismus (Epimorphismus).

Proposition 8.27 *Ein Gruppenhomomorphismus φ ist genau dann injektiv, wenn $Ker\, \varphi = e$.*

Beweis. Ist $\varphi(x_1) = \varphi(x_2)$, so ist

$$\varphi(x_1 \circ x_2^{-1}) = \varphi(x_1) \circ (\varphi(x_2))^{-1} = e'.$$

Deshalb ist $x_1 \circ x_2^{-1} = e$ (d.h. $x_1 = x_2$), falls $Ker\, \varphi = e$.
Andererseits, ist φ injektiv, so ist $Ker\, \varphi = e$ (weil $\varphi(e) = e'$ gilt) □

8.2 Der Satz von Lagrange

Definition 8.28 Sei X eine Menge, $X \times X = \{(x, y) \ : \ x, y \in X\}$ die Menge aller Paare. Eine beliebige Teilmenge $R \subset X \times X$ heißt eine **Relation auf** X.

Bezeichnung 8.29 Sei $R \subset X \times X$ eine Relation auf X. Es ist üblich kurz $x \sim_R y$ (oder einfach $x \sim y$) statt $(x, y) \in R$ zu schreiben.

Definition 8.30 Sei X eine Menge. Eine Relation $R \subset X \times X$ heißt eine **Äquivalenzrelation auf** X, wenn die folgende Bedingungen erfüllt sind:
(i) *Reflexivität*: für alle $x \in X$ gilt $x \sim_R x$;
(ii) *Symmetrie*: ist $x \sim_R y$, so ist $y \sim_R x$;
(iii) *Transitivität* : ist $x \sim_R y$ und $y \sim_R z$, so ist $x \sim_R z$

Definition 8.31 Sei R eine Äquivalenzrelation auf X, $x \in X$ ein beliebiges Element. Dann heißt die Teilmenge

$$R(x) = \{y \in X \ : \ y \sim_R x\} \subset X$$

Äquivalenzklasse von x bezüglich R, und x heißt **Vertreter von** $R(x)$ (wegen 8.30(i) gilt $x \in R(x)$).

Proposition 8.32 *Sei R eine Äquivalenzrelation auf X. Dann gilt für alle $x, y \in X$ entweder $R(x) = R(y)$, oder $R(x) \cap R(y) = \emptyset$.*

Beweis. Angenommen $R(x) \cap R(y) \neq \emptyset$. Sei $z \in R(x) \cap R(y)$. Dann gilt $z \sim_R x$ und $z \sim_R y$. Nach 8.30(ii), gilt auch $x \sim_R z$. Nach 8.30(iii) gilt $x \sim_R y$ und damit erhält man $R(y) \subset R(x)$. Andererseits, $y \sim_R (x)$ (8.30 (ii)). Deshalb hat man $R(x) \subset R(y)$ (8.30 (iii)). Also $R(x) = R(y)$. □

Definition 8.33 Sei $(G, \cdot)$ eine Gruppe, $H \subset G$ eine Untergruppe. Wir definieren die folgende Relation auf G:

$$x \sim_H y \Leftrightarrow \exists h \in H \ : \ y = xh.$$

Proposition 8.34 *Die Relation $x \sim_H y$ ist eine Äquivalenzrelation.*

Beweis. Reflexivität: Man hat $x \sim_H x$, weil H das neutrale Element e besitzt (s. 8.14) und $x = xe$ gilt.
Symmetrie: Ist $x \sim_H y$ (d.h. $\exists h \in H \ : \ y = xh$), so ist $yh^{-1} = xhh^{-1} = x$, d.h. $y \sim_H x$ (weil $h^{-1} \in H$!).
Transitivität: Ist $x \sim_H y$ und $y \sim_H z$, so existieren $h_1, h_2 \in H$ mit $y = xh_1$, $z = yh_2$ und damit $z = xh_1h_2$, d.h. $x \sim_H z$ (weil $h_1h_2 \in H$!). □

Definition 8.35 Die Äquivalenzklasse eines Elements $g \in G$ bzgl. der Äquivalenzrelation $\sim_H$ heißt **Linksnebenklasse** von g. Diese Klasse wird mit

$$gH := \{gh \ : \ h \in H\}$$

bezeichnet.

Definition 8.36 Hat eine Gruppe G nur endlich viele Linksnebenklassen bzgl. H, so heißt die Anzahl k der Linksnebenklassen der **Index von** H in G und wird mit $[G : H]$ bezeichnet.

Aus 8.32 und 8.34 folgt:

Korollar 8.37 *Sei $(G, \cdot)$ eine Gruppe, $H \subset G$ eine Untergruppe. Dann ist G eine disjunkte Vereinigung der Linksnebenklassen.*

Satz 8.38 (Satz von Lagrange) *Sei $(G, \cdot)$ eine endliche Gruppe, $H \subset G$ eine Untergruppe. Dann gilt*

$$|G| = [G : H]|H|.$$

Beweis. Sei

$$g_1H, \ldots, g_kH$$

die Gesamtheit aller Linksnebenklassen mit $g_iH \neq g_jH$ für $i \neq j$
(d.h. $k = [G : H]$). Man hat

$$|G| = \sum_{i=1}^{k} |g_iH|.$$

Jede Linksnebenklasse g_iH ist aber gleichmächtig zu H, weil die natürliche Abbildung

$$H \to g_iH$$
$$h \mapsto g_ih$$

bijektiv ist ($x \mapsto g_i^{-1}x$ ist die Umkehrabbildung $g_iH \to H$). Deswegen gilt

$$|G| = k|H| = [G : H]|H|.$$

□

Aus dem Satz von Lagrange folgt:

Korollar 8.39 *Sei G eine endliche Gruppe, $g \in G$ ein beliebiges Element. Dann ist* $\operatorname{Ord} g$ *ein Teiler von $|G|$.*

Beweis. Sei $\operatorname{Ord} g = n$ und

$$H :=< g >= \{e, g, g^2, \ldots, g^{n-1}\}$$

die endliche zyklische Untergruppe mit dem Erzeuger g. Dann gilt $n = |H|$. □

8.3 Ringe und Körper

Definition 8.40 Eine Menge R zusammen mit zwei Verknüpfungen

$$+ \ : \ R \times R \to R \ \ (a,b) \mapsto a+b$$

$$\cdot \ : \ R \times R \to R \ \ (a,b) \mapsto a \cdot b$$

heißt **Ring**, wenn folgendes gilt:
(1) Die Menge R zusammen mit Addition "+" is eine abelsche Gruppe;
(2) Die Multiplikation "$\cdot$" is assoziativ;
(3) Es gelten die **Distributivgesetze**:

$$a \cdot (b+c) = a \cdot b + a \cdot c, \quad \forall a,b,c,$$

$$(a+b) \cdot c = a \cdot c + b \cdot c \quad \forall a,b,c.$$

Bemerkung 8.41 Sei $0 \in R$ das Nullelement des Ringes R. Aus dem Distributivgesetz folgt sofort:

$$0 \cdot a = a \cdot 0 = 0.$$

Definition 8.42 Ein Ring heißt **kommutativ**, wenn $a \cdot b = b \cdot a \ \forall a,b \in R$. Ein Element $1 \in R$ heißt **Einselement**, wenn $1 \cdot a = a \cdot 1 = a$ für alle $a \in R$.

Beispiel 8.43 Die folgenden Menge mit üblichen Operationen sind Ringe:
(1) Die Mengen $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$;
(2) Die Mengen der quadratischen Matrizen $M(n \times n, K)$ mit Einträgen aus K, wobei $K = \mathbb{Z}, \mathbb{Q}$ oder $\mathbb{R}$.
(3) Die Menge der reellwertigen Funktionen $f : \mathbb{R} \to \mathbb{R}$.
(4) Ist R ein Ring, so bildet die Menge aller quadratischen Matrizen $M(n \times n, R)$ mit Einträgen aus R einen Ring.

Beispiel 8.44 In der albelschen Gruppe $\mathbb{Z}/m\mathbb{Z}$ der Restklassen modulo m kann man durch

$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$

eine Multiplikation erklären. Diese Definition des Multiplikation hängt nicht von der Auswahl der Representanten ab: ist $a - a' = mk$ und $b - b' = ml$, so folgt

$$a \cdot b = (a' + mk) \cdot (b' + ml) = a' \cdot b' + m(b'k + a'l + mkl),$$

d.h. $a \cdot b - a' \cdot b'$ is durch m teilbar.

Definition 8.45 Eine Menge K zusammen mit zwei Verknüpfungen

$$+ \ : \ K \times K \to K \ \ (a,b) \mapsto a+b$$

$$\cdot \ : \ K \times K \to K \ \ (a,b) \mapsto a \cdot b$$

heißt **Körper**, wenn folgendes gilt:
(1) K ist ein kommutativer Ring;
(2) Die Teilmenge $K^* := K \setminus \{0\}$ ist eine abelsche Gruppe.

Beispiel 8.46 Die Mengen $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ sind Körper.

Proposition 8.47 *Jeder Körper K besitzt mindestens zwei Elemente 0 und 1 ($0 \neq 1$). Es gelten die folgenden Rechenregeln:*
(1) $a \cdot b = 0 \ \Rightarrow a = 0$ oder $b = 0$;
(2) $a \cdot (-b) = -a \cdot b$ und $(-a) \cdot (-b) = ab$.
(3) $x \cdot a = y \cdot a \ \Rightarrow x = y$, falls $a \neq 0$.

Definition 8.48 Ein Ring R heißt **nullteilerfrei**, wenn für alle $a, b \in R$ aus $a \cdot b = 0$ stes $a = 0$ oder $b = 0$ folgt.

Bemerkung 8.49 Jeder Körper ist nullteilerfrei (s. 8.47 (1)).

Proposition 8.50 *Der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ ist nullteilerfrei genau dann, wenn m eine Primzahl ist.*

Beweis. Ist m keine Primzahl und $m = k \cdot l$ mit $1 < k, l < m$, so ist $\overline{k} \neq 0$ und $\overline{l} \neq 0$, aber $\overline{k} \cdot \overline{l} = \overline{m} = \overline{0}$. Sei m eine Primzahl und $\overline{a} \cdot \overline{b} = \overline{ab} = \overline{0}$. Dann teil m das Produkt ab. Da m Primzahl ist, muß m ein Primfaktor von a oder von b sein, damit gilt entweder $\overline{a} = 0$ oder $\overline{b} = 0$. □

Proposition 8.51 *Ein nullteilerfreier, kommutativer Ring K mit endlich vielen Elementen und Einselement ist ein Körper. Insbesondere ist $\mathbb{Z}/m\mathbb{Z}$ ein Körper, wenn m eine Primzahl ist.*

Beweis. Wir zeigen, dass für jedes $a \in K^*$ die Multiplikation

$$K^* \to K^*, \ \ x \mapsto a \cdot x$$

eine surjektive Abbildung ist. Weil K^* endlich ist, genügt es zu zeigen, dass diese Abbildung injektive ist:
Ist $a \cdot x = a \cdot x'$, so ist $0 = a \cdot x - a \cdot x' = a(x - x')$ und damit $x - x' = 0$ (weil $a \neq 0$). Aus der Sujektivität folgt, dass es ein b gibt, so dass $a \cdot b = 1$. Wegen der

Kommutativität erhalten wir auch $b \cdot a = 1$ Deswegen gibt es zu jedem $a \in K^*$ das inverse Element $a^{-1} = b$. □

Definition 8.52 Sei R ein Ring mit Einselement 1. Die **Charakteristik** von R ist definiert durch

$$\text{char}(R) := \begin{cases} 0, \ \text{falls} \, n \cdot 1 := \underbrace{1 + \cdots + 1}_{n-\text{mal}} \neq 0 \ \text{für alle} \ n \geq 1, \\ \min\{n \ : \ \underbrace{1 + \cdots + 1}_{n-\text{mal}} = 0, \ \text{sonst.} \end{cases}$$

Proposition 8.53 *Ist K ein Körper, so ist* $\text{char}(K)$ *etweder Null oder eine Primzahl.*

Beweis. Ist $n = char(K)$ keine Primzahl und $n = k \cdot l$ mit $1 < k, l < n$, so ist $k \cdot 1 \neq 0$ und $l \cdot 1 \neq 0$, aber $(k \cdot 1) \cdot (l \cdot 1) = n \cdot 1 = 0$. Widerspruch zu 8.47 (1). □

Definition 8.54 Ist R ein Ring und $R' \subset R$ eine Teilmenge, so heißt R' **Unterring**, wenn R' bzgl. der Addition Untergruppe ist und für alle $a, b \in R'$ gilt $a \cdot b \in R'$. Sind $(R, +, \cdot)$ und $(S, \oplus, *)$ zwei Ringe, so heißt eine Abbildung $\varphi \ : \ R \to S$ ein **Ringhomomorphismus** , wenn fúr alle $a, b \in R$ gilt:

$$\varphi(a + b) = \varphi(a) \oplus \varphi(b), ; \ \varphi(a \cdot b) = \varphi(a) * \varphi(b).$$

Ein bijektiver Ringhomomorphismus heißt **Isomorphismus** von Ringen R und S. Zwei Ringe R und S heißen **isomorph**, wenn es ein Ringisomorphismus gibt.

8.4 Komplexe Zahlen

Satz 8.55 *Die Menge $\mathbb{C}$ aller quadratischen Matrizen*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \ a, b \in \mathbb{R}$$

mit üblichen Matrizenmultiplikation und Matrizenaddition ist ein Körper.

Beweis. Ist $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ und $B = \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix}$, so ist

$$A \cdot B = \begin{pmatrix} aa' - bb' & -ab' - a'b \\ ab' + a'b & aa' - bb' \end{pmatrix} \in \mathbb{C}$$

und

$$A + B = \begin{pmatrix} a + a' & -b - b' \\ b + b' & a + a' \end{pmatrix} \in \mathbb{C}.$$

Die Multiplikation der Matrizen aus $\mathbb{C}$ is kommutativ. Es gelten Distributivgesetze. Die Matrix $\mathbf{0} := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ist Nullelement. Die Einheitsmatrix $\mathbf{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist Einselement. Zu jeder Matrix $\mathbf{0} \neq A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{C}$ gibt es das Inverse Matrix

$$A^{-1} = \frac{1}{a^2+b^2} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ \frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{pmatrix} \in \mathbb{C}.$$

Sei

$$\mathbf{I} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}.$$

Man kann jede Matrix A aus $\mathbb{C}$ in der Gestahlt

$$A = a \cdot \mathbf{1} + b \cdot \mathbf{I}$$

schreiben. Wir bemerken, dass $\mathbf{I}^2 = -\mathbf{1}$. □

Definition 8.56 Der Körper $\mathbb{C}$ im o.g Beispiel nennt man der Körper der **komplexen Zahlen**. Anstatt der Bezeichnungen $\mathbf{1}$ und $\mathbf{I}$ verwendet man einfach 1 und i (wobei $i^2 = -1$). Die Dastellung $c = a \cdot 1 + b \cdot i$ schreibt man einfach als $c = a + bi$, wobei $a := \operatorname{Re} c$ **Realteil** und $b := \operatorname{Im} c$ **Imaginärteil** von c genannt wird. Damit identifiziert man die Menge der reellen Zahlen $\mathbb{R}$ mit der komplexen Zahlen $c \in \mathbb{R}$, so dass $\operatorname{Im} c = 0$. Das Element $\overline{c} := a - bi \in \mathbb{C}$ heißt die zu c **konjugierte komplexe Zahl**. Die Zahl $|c| := \sqrt{a^2+b^2}$ heißt der **Absolutbetrag** von $c = a + bi$.

Proposition 8.57 *Für die komplexen Zahlen $z, z_1, z_2 \in \mathbb{C}$ gelten die folgenden Rechenregeln:*
(1) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$*;*
(2) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$
(3) $|z_1 + z_2| \leq |z_1| + |z_2|$ *(Dreiecksungleichung);*
(4) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$*;*

Definition 8.58 Eine Matrix

$$c = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

kann in der Gestalt

$$c = \sqrt{a^2+b^2} \begin{pmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{-b}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{pmatrix}$$

geschrieben werden. Wegen

$$-1 \leq \frac{a}{\sqrt{a^2+b^2}} \leq 1, \quad -1 \leq \frac{b}{\sqrt{a^2+b^2}} \leq 1$$

können wir setzen:

$$\cos\alpha := \frac{a}{\sqrt{a^2+b^2}}, \quad \sin\alpha := \frac{b}{\sqrt{a^2+b^2}}.$$

Dann erhalten wir:

$$c = \sqrt{a^2+b^2}\begin{pmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{pmatrix} = |c|e^{i\alpha},$$

wobei

$$e^{i\alpha} := \begin{pmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{pmatrix} = \cos\alpha + i\sin\alpha.$$

Die Zahl α nennt man das **Argument** von c:

$$c = |c| \cdot e^{i\arg c}.$$

Diese Darstellung einer komplexen Zahl c nennt man **trigonometrische Form** von c. (Das Argument $\arg c$ in der trigonometrischen Form ist eindeutig nur bis auf ein Vielfaches von 2π bestimmt.)

Proposition 8.59 *Es seine z_1, z_2 komplexe Zahlen. Dann gilt*

$$e^{i\arg z_1 \cdot z_2} = e^{i\arg z_1 + \arg z_2},$$

oder

$$\arg(z_1 \cdot z_2) - (\arg z_1 + \arg z_2) \in 2\pi\mathbb{Z}.$$

Beweis. Es sei $\alpha_1 = \arg z_1$, $\alpha_2 = \arg z_2$. Wir benutzen die bekannten Formeln

$$\cos(\alpha_1+\alpha_2) = \cos\alpha_1\cos\alpha_2 - \sin\alpha_1\sin\alpha_2,$$

$$\sin(\alpha_1+\alpha_2) = \cos\alpha_1\sin\alpha_2 + \sin\alpha_1\cos\alpha_2.$$

um zu zeigen, dass gilt

$$(\cos\alpha_1 + i\sin\alpha_1)(\cos\alpha_2 + i\sin\alpha_2) = \cos(\alpha_1+\alpha_2) + i\sin(\alpha_1+\alpha_2).$$

□

Satz 8.60 *Sei $z = \rho e^{i\alpha} = \rho(\cos\alpha + i\sin\alpha)$ eine komplexe Zahl und n eine positive ganze Zahl. Dann besitzt die Gleichung*

$$x^n = z$$

genau n Lösungen

$$x_k = \sqrt[n]{\rho} e^{i\beta_k}, \quad k = 0, 1, 2, \ldots, n-1,$$

wobei

$$\beta_k = \frac{\alpha}{n} + \frac{2\pi k}{n}.$$

Beweis. Ist $x = re^{i\beta}$, so ist $x^n = r^n e^{in\beta}$. Aus der Gleichung

$$x^n = r^n e^{in\beta} = \rho e^{i\alpha}$$

folgt, dass $\rho = r^n$ und

$$n\beta = \alpha + 2\pi k, \quad k \in \mathbb{Z}.$$

Damit gilt

$$r = \sqrt[n]{\rho}, \quad \beta = \frac{\alpha}{n} + \frac{2\pi k}{n}, \quad k = 0, 1, \ldots, n-1.$$

□

8.5 Polynome

Definition 8.61 Sei K ein Körper und t eine Unbestimmte. Den formallen Ausdruck der Gestalt

$$f(t) = a_0 + a_1 t + \cdots + a_n t^n = \sum_i a_i t^i,$$

wobei $a_0, a_1, \ldots, a_n \in K$ nennen wir **Polynom mit Koeffizienten in** K. Die Menge all solcher Polynome bezeichnen wir mit $K[t]$. Oft schreiben wir f statt $f(t)$. Sind alle Koeffizienten a_i von $f(t)$ gleich Null, so heißt $f(t)$ **Nullpolynom** (Man schreibt $f = 0$. **Der Grad** von f ist erklärt als

$$\deg f := \begin{cases} -\infty, & \text{falls } f = 0, \\ \max\{k \in \mathbb{N} \,:\, a_k \neq 0\}, & \text{sonst.} \end{cases}$$

Das Polynom f vom Grad n heißt **normiert**, wenn $a_n = 1$.

Definition 8.62 Die Menge $K[t]$ besitzt **Addition** und **Multiplikation**: Ist $f = a_0 + a_1 t + \cdots + a_n t^n = \sum_{i\geq 0} a_i t^i$ und $g = b_0 + b_1 t + \cdots + b_m t^m = \sum_{i\geq 0} b_i t^i$ (wir setzen $a_{n+k} = b_{m+k} = 0$ für $k \geq 1$, so ist

$$f + g := \sum_{i\geq 0} (a_i + b_i) t^i,$$

$$f \cdot g = c_0 + c_1 t + \cdots + c_{n+m} t^{n+m} = \sum_{i \geq 0} c_i t^i,$$

wobei

$$c_i = \sum_{k+l=i} a_k b_l.$$

Insbesondere ist

$$c_0 = a_0 b_0$$

$$c_1 = a_0 b_1 + a_1 b_0$$

$$c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$$

$$\cdots$$

$$c_{n+m} = a_n b_m.$$

Ist $f \cdot g = h$, so nennt man f und **Teiler** von h und schreibt $f|h$ (man sagt auch, dass h durch f **teilbar** ist.

Proposition 8.63 *Die Menge $K[t]$ ist ein kommutativer Ring mit 1. $K[t]$ ist nullteilerfrei und gilt*

$$\deg(f \cdot g) = \deg f + \deg g.$$

(Es wird vorausgesetzt, dass $n + (-\infty) = m + (-\infty) = -\infty + (-\infty) = -\infty$).
Der Ring $K[t]$ wird **Polynomring über** *K genannt.*

Satz 8.64 *Sind $f, g \in K[t]$ Polynome, und ist $g \neq 0$, so gibt es dazu eindeutig bestimmte Polynome $q, r \in K[t]$ derart, dass*

$$f = q \cdot g + r \;\; \textit{und} \;\; \deg r < \deg g.$$

Definition 8.65 Die Darstellung von f als $q \cdot g + r$ mit $\deg r < \deg g$ nennt man **Division von f durch g mit Rest**. Das Polynom q heißt **Quotient** und das Polynom r **Rest**.

Definition 8.66 Sei $f, g \in K[t]$ zwei Polynome. Ein Polynom $p \in K[t]$ heißt der **größte gemeinsame Teiler** von f und g, wenn gilt
(1) $p|f$ und $p|g$;
(2) ist q ein anderes Polynom mit $q|f$ und $q|g$, so gilt $q|p$. Der größte gemeinsame Teiler von f und g wird mit $g.g.T.(f, g)$ bezeichnet.

Der größte gemeinsame Teiler zweier Polynome $f, g \in K[t]$ immer existiert und bis auf Multiplikation mit einer Konstante eindeutig bestimmt. Man kann $g.g.T.(f, g)$ durch den eulidischen Algorithmus wie folgt finden:

Satz 8.67 (Eulidischer Algorithmus) *Seien $f_1, f_2 \in K[t]$ zwei Polynome mit* $\mathrm{degf}_1 \geq \mathrm{degf}_2$. *Man bildet eine endliche Folge der Polynome*

$$f_1 = q_1 f_2 + f_3,$$

$$f_2 = q_2 f_3 + f_4,$$

$$f_3 = q_3 f_4 + f_5,$$

$$\dots$$

$$f_{r-3} = q_{r-3} f_{r-2} + f_{r-1},$$

$$f_{r-2} = q_{r-2} f_{r-1} + f_r,$$

$$f_{r-1} = q_{r-1} f_r,$$

wobei $\mathrm{deg}f_3 > \mathrm{deg}f_4 > \mathrm{deg}f_5 > \cdots > \mathrm{deg}f_r \geq 0$ $(f_r \neq 0)$. *Dann gilt*

$$g.g.T.(f_1, f_2) = f_r.$$

Lemma 8.68 *Seien $f, g, h \in K[t]$ und $f|g$, $f|h$. Dann gilt*
(1) *$f|(g+h)$ und $f|(g-h)$;*
(2) *$f|qg$ für alle $q \in K[t]$;*

Beweis Ist $g = fq_1,\ h = fq_2$, so ist

$$(g \pm h) = fq_1 \pm fq_2 = f(q_1 \pm q_2),$$

$$g \cdot q = fq_1 q = f(q_1 q).$$

□

Beweis des Satzes 8.67. Zu nächst zeigen wir, dass f_r ein gemeinsamer Teiler von f_1 und f_2 ist. Dafür bemerken wir, dass aus

$$f_{r-1} = q_{r-1} f_r$$

und

$$f_{r-2} = q_{r-2} f_{r-1} + f_r$$

die Teilbarkeit von f_{r-1} und f_{r-2} durch f_r folgt (s. 8.68). Weiter folgt aus

$$f_{r-3} = q_{r-3} f_{r-2} + f_{r-1},$$

dass auch f_{r-3} durch f_r teilbar ist. Mit gleichen Überlegungen erhalten wir, dass alle Polynome $f_{r-1}, f_{r-2}, \ldots, f_2, f_1$ durch f_r teilbar sind. Also ist f_r ein gemeinsamer Teiler von f_1 und f_2.
Nun sei $q \in K[t]$ ein anderer gemeinsamer Teiler von f_1 und f_2. Aus

$$f_1 = q_1 f_2 + f_3$$

folgt, dass

$$f_3 = f_1 - q_1 f_2.$$

Wegen 8.68 erhalten wir, dass f_3 durch q teilbar ist. Weiterhin folgt aus

$$f_4 = f_2 - q_2 f_3,$$

dass f_4 durch q teilbar ist. Mit gleichen Überlegungen erhalten wir, dass alle Polynome $f_1, f_2, f_3, \ldots, f_r$ durch q teilbar sind. Insbesondere ist f_r durch q teilbar. □

Korollar 8.69 *Für den euklidischen Algorithmus wie oben gilt*

$$g.g.T.(f_1, f_2) = g.g.T.(f_2, f_3) = \cdots = g.g.T.(f_{r-2}, f_{r-1}) = f_r.$$

Korollar 8.70 *Seien f, g Polynome aus $K[t]$. Ist $h = g.g.T.(f, g)$, so gibt es Polynome $u, v \in K[t]$ mit*

$$h = uf + vg.$$

Definition 8.71 Sei K ein Körper und $f = a_0 + a_1 t + \cdots + a_n t^n \in K[t]$ ein Polynom. Ein Element $\lambda \in K$ heißt **Nullstelle** von f, wenn gilt

$$f(\lambda) := a_0 + a_1\lambda + \cdots + a_n\lambda^n = 0.$$

Proposition 8.72 *Sei $\lambda \in K$. Ein Polynom $f \in K[t]$ ist durch $g = (t - \lambda)$ teilbar genau dann, wenn λ eine Nullstelle von f ist.*

Beweis. Wir schreiben $f(t) = q(t)g(t) + r(t)$ mit $\deg r < \deg g = 1$. Wegen $\deg r \leq 0$ erhalten wir, dass $r = 0 \Leftrightarrow f(\lambda) = 0$. □

Definition 8.73 Ein Element $\lambda \in K$ heißt **Nullstelle** des Polynoms $f \in K[t]$ der **Vielfachheit** k, wenn gilt
(1) $(t - \lambda)^k | f$;
(2) $(t - \lambda)^{k+1}$ kein Teiler von f.

Der folgende Satz wurde von C. F. Gauß ertsmals 1799 bewiesen:

Satz 8.74 (Fundamentalsatz der Algebra) *Jedes Polynom $f \in \mathbb{C}[t]$ mit $\deg f > 0$ hat mindestens eine Nullstelle $\lambda \in \mathbb{C}$.*

Korollar 8.75 *Jedes Polynom $f \in \mathbb{C}[t]$ zerfällt in Linearfaktoren, d.h. es gibt eine komplexe Zahl a und komplexe Zahlen $\lambda_1, \ldots, \lambda_n$ mit $n = \deg f$, so dass*

$$f = a(t - \lambda_1) \cdots (t - \lambda_n).$$

Satz 8.76 *Jedes Polynom $f \in \mathbb{R}[t]$ mit $\deg f = n > 0$ besitzt eine Zerlegung*

$$f = a(t - \lambda_1) \cdots (t - \lambda_r) \cdot g_1 \cdots g_m,$$

wobei $a, \lambda_1, \ldots, \lambda_r$ reelle Zahlen sind $(a \neq 0)$ und $g_1, \ldots, g_m \in \mathbb{R}[t]$ normierte Polynome vom Grad 2.

Lemma 8.77 *Ist $f \in \mathbb{R}[t]$ und $\lambda \in \mathbb{C}$ eine Nullstelle von f, so ist $\overline{\lambda}$ auch eine Nullstelle von f.*

Beweis Ist $f(t) = a_0 + a_1 t + \cdots + a_n t^n$, so ist

$$f(\lambda) = a_0 + a_1 \lambda + \cdots + a_n \lambda^n = 0$$

und

$$0 = \overline{0} = \overline{a_0 + a_1\lambda + \cdots + a_n\lambda^n} = \overline{a_0} + \overline{a_1\lambda} + \cdots + \overline{a_n\lambda^n} = a_0 + a_1\overline{\lambda} + \cdots + a_n\overline{\lambda}^n = f(\overline{\lambda})$$

□

Beweis des Satzes 8.76. Wir führen Induktion über den Grad n von f.
Für $n = 1$ ist f ein lineares Polynom: $f = at + b = a(t - \lambda_1)$, wobei $\lambda_1 = -b/a$.
Sei $n > 1$. Besitzt f eine Nullstelle $\lambda \in \mathbb{R}$, so ist f durch $(t - \lambda)$ teilbar: $f = (t - \lambda)g$, wobei $g \in \mathbb{R}[t]$ und $\deg g = n - 1$. Nach Induktionsannahme gibt es eine Zerlegung

$$g = a(t - \lambda_1) \cdots (t - \lambda_r) \cdot g_1 \cdots g_m$$

und damit

$$f = a(t - \lambda)(t - \lambda_1) \cdots (t - \lambda_r) \cdot g_1 \cdots g_m.$$

Besitzt f eine komplexe (nicht reeelle) Nulstelle $\lambda = x + yi \in \mathbb{C}$, $\lambda \notin \mathbb{R}$, so ist die konjudierte Zahl $\overline{\lambda} = x - yi$ auch eine Nullstelle von f (s. Lemma 8.77). Wir bezeichnen

$$h := (t - \lambda)(t - \overline{\lambda}) = t^2 - 2xt + (x^2 + y^2).$$

Durch Division mit Rest erhalten wir

$$f = q \cdot h + r,$$

wobei $\deg r \leq 1$. Wir haben

$$0 = f(\lambda) = q(\lambda) \cdot h(\lambda) + r(\lambda) = q(\lambda) \cdot 0 + r(\lambda) = r(\lambda).$$

Wegen $r \in \mathbb{R}[t]$ und $\deg r \leq 1$ die Gleichung $r(\lambda) = 0$ ist nur dann möglich, wenn r ein Nullpolynom ist. Also gilt

$$f = q \cdot h.$$

Nach Induktionsannahme gibt es eine Zerlegung

$$q = a(t - \lambda_1) \cdots (t - \lambda_r) \cdot g_1 \cdots g_m.$$

Damit erhalten wir die gewünschte Zerlegung von f:

$$f = a(t - \lambda_1) \cdots (t - \lambda_r) \cdot h g_1 \cdots g_m.$$

□

Korollar 8.78 *Jedes Polynom $f \in \mathbb{R}[t]$ von ungeradem Grad hat mindestens eine reelle Nullstelle.*

9 Vektorräume und lineare Abbildungen

9.1 Vektorräume über einem Körper K

Definition 9.1 Sei K ein beliebiger Körper (z. B. $\mathbb{Q}, \mathbb{R}, \mathbb{C}, oder \mathbb{Z}/p\mathbb{Z}$ mit eine Primzahl p). Eine abelsche Gruppe $(V, +)$ mit einer zusätzlichen Verknüpfung

$$K \times V \to V,$$

$$(\lambda, v) \mapsto \lambda v, \quad (\lambda \in K,\ v \in V)$$

heißt **Vektorraum über** K (oder K**-Vektorraum**), wenn die folgenden Bedingungen erfüllt sind:
(1) $(\lambda \cdot \mu)v = (\lambda(\mu v)), \ \forall \lambda, \mu \in K, \ \forall v \in V$;
(2) $(\lambda + \mu)v = \lambda v + \mu v, \ \forall \lambda, \mu \in K, \ \forall v \in V$;
(3) $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 \ \forall \lambda \in K, \ \forall v_1, v_2 \in V$;
(4) $1v = v, \ \forall v \in V$.
Elemente eines K-Vektorraums V heißen **Vektoren**.

Beispiel 9.2 Sei K ein beliebiger Körper. Die Menge

$$V = K^n := \{(x_1, \ldots, x_n) \ : \ x_i \in K \ \ (1 \leq i \leq n)\}$$

ist ein K-Vektorraum bzgl. der Operationen:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

$$\lambda(x_1, \ldots, x_n) := (\lambda \cdot x_1, \ldots, \lambda \cdot x_n).$$

Definition 9.3 Sei V ein K-Vektorraum und $\{v_1, \ldots, v_k\}$ eine Menge von Vektoren in V. Man sagt, dass ein Vektor $v \in V$ als **Linearkombination von** $v_1, \ldots, v_k$ **darstellbar**, wenn gilt

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

mit einigen Elementen $\lambda_1, \ldots, \lambda_k$ aus K. Die Elemente $\lambda_1, \ldots, \lambda_k$ aus K nennt man **Koeffizienten der Linearkombination**. Sind alle $\lambda_1, \ldots, \lambda_k$ gleich $0 \in K$, so nennt man $\lambda_1 v_1 + \cdots + \lambda_k v_k$ die **triviale Linearkombination**.

Definition 9.4 Die Vektoren $v_1, \ldots, v_k \in V$ heißen **linear unabhängig**, wenn der Nullvektor $\mathbf{0} \in V$ nur als triviale Linearkombination von $v_1, \ldots, v_k$ darstellbar ist, d.h.,

$$\mathbf{0} = \lambda_1 v_1 + \cdots + \lambda_k v_k \ \Rightarrow \lambda_1 = \cdots = \lambda_k = 0.$$

Sind Vektoren $v_1, \ldots, v_k$ nicht linear unabhängig, so heißen $v_1, \ldots, v_k$ **linear abhängig**. Die lineare Abhängigkeit von $v_1, \ldots, v_k$ bedeutet, dass es eine nicht-triviale Linearekombination $\lambda_1 v_1 + \cdots + \lambda_k v_k$ gibt (mindestens für ein $\lambda_i \in K$ gilt $\lambda_i \neq 0$), die gleich dem Nullvektor ist.

Definition 9.5 Sei V ein K-Vektorraum und $\{v_i\}_{i\in I} \subseteq V$ eine Teilmenge (diese Teilmenge kann unendlich sein). Die Teilmenge $\{v_i\}_{i\in I}$ heißt eine **Basis von** V, wenn gilt
(1) jede endliche Teilmenge $\{v_1, \ldots, v_k\} \subseteq \{v_i\}_{i\in I}$ ist linear unabhängig;
(2) jeder Vektor $v \in V$ ist als Linearkombination von endlich vielen Vektoren aus $\{v_i\}_{i\in I}$ darstellbar.

Proposition 9.6 *Sei V ein K-Vektorraum. Eine Teilmenge $M = \{v_i\}_{i\in I} \subseteq V$ ist Basis genau dann, wenn jede endliche Teilmenge in M linear unabhängig ist, und es keine echt größere Teilmenge $M \subset M'$ ($M \neq M'$) gibt, so dass jede endliche Teilmenge in M' auch linear unabhängig ist.*

Beweis. Eine einfache Übung.

Satz 9.7 *Jeder Vektorraum V besitzt eine Basis. Je zwei Basen von V haben gleiche Mächtigkeit.*

Der Beweis dieses Satzes für Basen beliebiger Mächtigkeit brauch das sog. **Lemma von Zorn**. Besitzt V eine endliche Basis, so kann man wie für Untervektorräume in $\mathbb{R}^n$ zeigen, dass alle Basen von V geliche Mächtigkeit besitzten.

Definition 9.8 Die Mächtigkeit einer Basis von K heißt **Dimension** von V über K und wird mit $\dim_K V$ bezeichnet. Ist $\dim_K V < \infty$, so sagt man, dass V ein K-Vektorraum **endlicher Dimension** ist.

Beispiel 9.9 Die Menge $V = K^n$ ist ein K-Vektorraum der Dimension n über k. Die Vektoren

$$v_1 = (1, 0, \ldots, 0),\ v_2 = (1, 0, \ldots, 0), \ldots, v_n = (0, 0, \ldots, 1)$$

bilden eine Basis von K^n.

Beispiel 9.10 Der Plynomring $K[t]$ ist ein K-Vektorraum. Die Teilmenge

$$\{1, t, t^2, \ldots, t^n, \ldots\} = \{t^i\}_{i\in\mathbb{Z}_{\geq 0}}$$

ist eine Basis von $K[t]$. Damit hat $K[t]$ unendliche Dimension über K.

Beispiel 9.11 $\mathbb{C}$ ist ein 2-dimensionaler $\mathbb{R}$-Vektorraum: $\dim_{\mathbb{R}}\mathbb{C} = 2$. Andererseits gilt $\dim_{\mathbb{C}}\mathbb{C} = 1$.

Beispiel 9.12 Die Menge $M(n \times m; K)$ von Matrizen ist ein K-Vektorraum der Dimension nm. Eine Basis von $M(n \times m; K)$ besteht aus Matrizen E_{ij} $(1 \leq i \leq m, 1 \leq j \leq n)$, wobei die Einträge a_{kl} von E_{ij} sind gleich $0 \in K$, wenn $k \neq i$ und $l \neq j$, und $a_{ij} = 1 \in K$.

Aus 9.6 folgt sofort:

Proposition 9.13 *Sei V ein K-Vektorrraum mit $\dim_K V = d$. Eine Teilmenge $\{v_1, \ldots, v_k\} \subset V$ ist eine Basis von V genau dann, wenn $k = d$ und die Vektoren $v_1, \ldots, v_k$ linear unabhängig sind.*

Definition 9.14 Sei V ein K-Vektorrraum. Eine Teilmenge $U \subset V$ heißt **K-Vektorunterraum**, wenn die folgenden Bedingungen erfüllt sind:
(1) $\forall x, y \in U \;\Rightarrow\; x + y \in U$;
(2) $\forall x \in U,\ \forall \lambda \in K \Rightarrow \lambda x \in U$.

Satz 9.15 *Sei V ein K-Vektorrraum mit $\dim_K V = d$ und $U \subset V$ ein K-Vektorunterraum. Dann gilt:*
(1) $k := \dim_K U \leq \dim_K V = d$;
(2) jede Basis $\{v_1, \ldots, v_k\}$ von U läßt sich bis auf eine Basis $\{v_1, \ldots, v_k, \ldots, v_d\}$ fortsetzen.
(3) $U = V$ genau dann, wenn $\dim_K U = \dim_K V$.

Definition 9.16 Seien V und W zwei K-Vektorräume. Eine Abbildung $f : V \to W$ heißt **linear**, wenn die folgenden Bedingungen erfüllt sind:
(1) $\forall x, y \in V$ gilt $f(x + y) = f(x) + f(y)$;
(2) $\forall x \in V$ und $\forall \lambda \in K$ gilt $f(\lambda x) = \lambda f(x)$.
Zwei K-Vektorräume V und W heißen **isomorph**, wenn es eine bijektive lineare Abbildung $f : V \to W$ gibt.

Satz 9.17 *Jeder K-Vektorraum V der Dimension d ist zum K-Vektorraum K^d isomorph.*

Beweis. Es sei $e_1, \ldots, e_d$ eine Basis von V. Dann definieren wir die Abbildung $f : K^d \to V$ als

$$f(x_1, \ldots, x_d) := x_1 e_1 + \cdots + x_d e_d, \;\; \forall\, (x_1, \ldots, x_d) \in K^d.$$

Offensichtlich ist f linear. Der Kern von f ist trivial (damit f ist injektiv), weil $e_1, \ldots, e_d$ linear unabhängig sind. Andererseits ist f surjektiv, da jeder Vektor aus V als Linearkombination von $e_1, \ldots, e_d$ darstellbar ist. Also ist f ein Isomorphismus. □

Satz 9.18 (Dimensionsformel für lineare Abbidungen) *Seien V und W zwei K-Vektorräume. Ist $f : V \to W$ eine lineare Abbildung und*

$$Ker\, f := \{v \in V \ : \ f(v) = 0\},$$

$$Im\, f := \{w \in W \ : \ \exists v \in V \text{ mit } f(v) = w\},$$

so gilt

$$\dim Im\, f + \dim Ker\, f = \dim(V).$$

Beweis. Es sei $v_1, \ldots, v_k$ eine Basis von $Ker\, f$, d.h. $k = \dim Ker\, f$. Wir setzen $v_1, \ldots, v_k$ bis auf einer Basis $v_1, \ldots, v_n$ fort. Nun zeigen wir, dass

$$f(v_{k+1}), \ldots, f(v_n)$$

eine Basis von $Im\, f$ ist.
Ist $v = x_1v_1 + \cdots x_nv_n$ ein Vektor aus V, so ist

$$f(v) = f(x_1v_1 + \cdots x_nv_n) = x_1f(v_1) + \cdots x_nf(v_n) = x_1f(v_{k+1}) + \cdots x_nf(v_n)$$

wegen $f(v_1) = \cdots = f(v_k) = 0$. Also ist $f(v)$ als Linearkombination von $f(v_{k+1}), \ldots, f(v_n)$ darstellbar.
Ist $\lambda_{k+1}f(v_{k+1}) + \cdots + \lambda_nf(v_n) = 0$, so ist $v := \lambda_{k+1}v_{k+1} + \cdots + \lambda_nv_n \in Ker\, f$. Deswegen gibt es eine Darstellung von v als Linearkombination

$$v = \lambda_1v_1 + \cdots + \lambda_kv_k.$$

Damit erhalten wir:

$$0 = v - v = \lambda_1v_1 + \cdots + \lambda_kv_k - \lambda_{k+1}v_{k+1} - \cdots - \lambda_nv_n.$$

Aus der linearen Unabhängigkeit von $v_1, \ldots, v_n$ folgt, dass alle λ_i $(i = 1, \ldots, n)$ gleich 0 sind. Insbesondere $\lambda_{k+1} = \cdots = \lambda_n = 0$. Also sind $f(v_{k+1}), \ldots, f(v_n)$ linear unabhängig und damit eine Basis von $Im\, f$. □

9.2 Matrizen von linearen Abbildungen

Satz 9.19 *Es sei $v_1, \ldots, v_n$ eine Basis von V. Eine lineare Abbildung f von V nach W ist durch Angabe von $f(v_1), \ldots, f(v_n) \in W$ eindeutig bestimmt.*

Beweis. Es sei $v \in V$ ein beliebiger Vektor. Dann gibt es eindeutig bestimmte $x_1, \ldots, x_n \in K$ mit

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Damit folgt

$$f(v) = f(x_1 v_1 + \cdots + x_n v_n) = f(x_1 v_1) + \cdots + f(x_n v_n) = x_1 f(v_1) + \cdots + x_n f(v_n).$$

Also ist $f(v)$ eindeutig bestimmt. □

Definition 9.20 Seien V und W zwei K-Vektorräume, $\{v_1, \ldots, v_n\}$ eine Basis von V, $\{w_1, \ldots, w_m\}$ eine Basis von W. Ist $f : V \to W$ eine lineare Abbildung, so schreiben wir

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

Die Koeffizienten (a_{ij}) bilden eine $m \times n$-Matrix A mit Einträgen aus K. Die Matrix $A = (a_{ij})$ heißt **Matrix der linearen Abbildung f in den Basen** $\{v_1, \ldots, v_n\}$ und $\{w_1, \ldots, w_m\}$.
Ist $V = W$ und $\{v_1, \ldots, v_n\} = \{w_1, \ldots, w_m\}$, so heißt die quadratische Matrix $A = (a_{ij}) \in M(n \times n; K)$ **Matrix der linearen Abbildung $f : V \to V$ in der Basis $\{v_1, \ldots, v_n\}$.**

Aus 9.19 folgt sofort:

Korollar 9.21 *Zwei lineare Abbildungen f und f' von V nach W sind gleich genau dann, wenn ihre Matrizen A und A' in den Basen $\{v_1, \ldots, v_n\} \subset V$ und $\{w_1, \ldots, w_m\} \subset W$ gleich sind.*

Bemerkung 9.22 Sei $f : V \to W$ eine lineare Abbildung, $v_1, \ldots, v_n$ eine Basis von V, $w_1, \ldots, w_m$ eine Basis von W. Dann ist Matrix A von f in den Basen $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ durch die folgende Matrizengleichung definiert:

$$(f(v_1), \ldots, f(v_n)) = (w_1, \ldots, w_n)A.$$

Es sei $f : V \to V$ eine lineare Abbildung (Endomorphismus), und sei $v_1, \ldots, v_n$ eine Basis von V. Dann ist Matrix A von f in der Basis $v_1, \ldots, v_n$ durch die folgende Matrizengleichung definiert:

$$(f(v_1), \ldots, f(v_n)) = (v_1, \ldots, v_n)A.$$

Proposition 9.23 *Es sei $f : V \to W$ eine lineare Abbildung. Mit $v_1, \ldots, v_n$ (bzw. $w_1, \ldots, w_m$) bezeichnen wir eine Basis von V (bzw. von W). Seien $x_1, \ldots, x_n$ die Koordinaten eines Vektors $x \in V$ in der Basis $v_1, \ldots, v_n$ und $y_1, \ldots, y_m$ die Koordinaten seines Bildes $y = f(x) \in W$ in der Basis $w_1, \ldots, w_m$. Dann gilt*

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Lemma 9.24 *Es seien $v_1, \ldots, v_n$ linear unabhängige Vektoren und $M, M' \in Mat(n \times m, K)$ zwei $(n \times m)$-Matrizen. Ist*

$$(v_1, \ldots, v_n)M = (v_1, \ldots, v_n)M',$$

so ist $M = M'$.

Beweis des Lemma. Aus

$$(v_1, \ldots, v_n)M = (v_1, \ldots, v_n)M'$$

folgt, dass

$$(v_1, \ldots, v_n)(M - M') = 0.$$

Da $v_1, \ldots, v_n$ linear unabhängig sind, erhalten wir $M - M' = 0$. □

Beweis der Proposition. Wir haben

$$x = x_1 v_1 + \cdots + x_n v_n = (v_1, \ldots, v_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Wegen der Linearität von f erhalten wir:

$$y = f(x) = x_1 f(v_1) + \cdots + x_n f(v_n) = (f(v_1), \ldots, f(v_n)) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Andererseits gilt

$$y = y_1 w_1 + \cdots + y_m w_m = (w_1, \ldots, w_m) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Durch Einsetzen in die vorletzte Gleichung erhalten wir:

$$(w_1, \ldots, w_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = (w_1, \ldots, w_n) A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Aus der linearen Unabhängigkeit von $w_1, \ldots, w_n$ und Lemma 9.24 folgt:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

□

Aus 9.18 folgt:

Korollar 9.25 *Es gilt*

$$\dim Ker\, f = n - \operatorname{Rang} A,\ \dim Im\, f = \operatorname{Rang} A.$$

Definition 9.26 Es seien $v_1, \ldots, v_n$ und $v'_1, \ldots, v'_n$ zwei Basen von V. Wir schreiben

$$v'_j = \sum_{i=1}^{n} t_{ij} v_i$$

Die Koeffizienten (t_{ij}) bilden eine $n \times n$-Matrix T mit Einträgen aus K. Die Matrix $T = (a_{ij})$ heißt **Basiswechselmatrix (Übergangsmatrix) von** $\{v_1, \ldots, v_n\}$ **zu** $\{v'_1, \ldots, v'_n\}$. Es gilt für die Basiswechselmatrix T

$$(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n) T.$$

Proposition 9.27 *Es sei T' die Basiswechselmatrix von $\{v'_1, \ldots, v'_n\}$ zu $\{v_1, \ldots, v_n\}$, d.h.*

$$(v_1, \ldots, v_n) = (v'_1, \ldots, v'_n) T'.$$

Dann ist T' die zu T inverse Matrix. Sind $x_1, \ldots, x_n$ (bzw. $x'_1, \ldots, x'_n$) Koordinaten eines Vektors v in der Basis $v_1, \ldots, v_n$ (bzw. in der Basis $v'_1, \ldots, v'_n$), so ist

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = T' \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Beweis. Aus $(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n)T$ und $(v_1, \ldots, v_n) = (v'_1, \ldots, v'_n)T'$ folgt, dass

$$(v'_1, \ldots, v'_n)E_n = (v'_1, \ldots, v'_n) = (v'_1, \ldots, v'_n)T'T$$

und

$$(v_1, \ldots, v_n)E_n = (v_1, \ldots, v_n) = (v_1, \ldots, v_n)TT'.$$

Da $v_1, \ldots, v_n$ und $v'_1, \ldots, v'_n$ linear unabhängig sind, erhalten wir wegen 9.24

$$E_n = T'T,\ E_n = TT'.$$

Also sind T und T' zueinander inverse Matrizen. □

Proposition 9.28 *Es seien $v_1, \ldots, v_n$, $v'_1, \ldots, v'_n$ zwei Basen von V und $w_1, \ldots, w_m$, $w'_1, \ldots, w'_m$ zwei Basen von W. Wir bezeichnen mit A (bzw. mit A') die Matrix einer linearen Abbildung $f : V \to W$ in den Basen $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ (bzw. in den Basen $v'_1, \ldots, v'_n$ und $w'_1, \ldots, w'_m$). Dann gilt*

$$A' = S^{-1}AT,$$

wobei T (bzw. S) die Basiswechselmatrix von $\{v_1, \ldots, v_n\}$ zu $\{v'_1, \ldots, v'_n\}$ (bzw. $\{w_1, \ldots, w_m\}$ zu $\{w'_1, \ldots, w'_m\}$) ist.

Beweis. Wir haben

$$(f(v'_1), \ldots, f(v'_n)) = (w'_1, \ldots, w'_m)A'.$$

Wegen der Linearität von f aus

$$(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n)T$$

erhalten wir:

$$(f(v'_1), \ldots, f(v'_n)) = (f(v_1), \ldots, f(v_n))T.$$

Also gilt

$$(f(v_1), \ldots, f(v_n))T = (w'_1, \ldots, w'_m)A' = (w_1, \ldots, w_n)SA'$$

und damit

$$(w_1, \ldots, w_m)AT = (w_1, \ldots, w_n)SA'.$$

Aus linearen Unabhängigkeit von $v_1, \ldots, v_n$ und Lemma 9.24 folgt:

$$AT = SA'.$$

Deswegen gilt

$$A' = S^{-1}AT.$$

□

Korollar 9.29 *Ist* $V = W$ *und* $\{v_1, \ldots, v_n\} = \{w_1, \ldots, w_m\}$, $\{v'_1, \ldots, v'_n\} = \{w'_1, \ldots, w'_m\}$, *so gilt*

$$A' = T^{-1}AT,$$

wobei T *die Basiswechselmatrix von* $\{v_1, \ldots, v_n\}$ *zu* $\{v'_1, \ldots, v'_n\}$ *ist.*

Proposition 9.30 *Es sei* $f : V \to W$ *eine lineare Abbildung und* A *die Matrix von* f *in den Basen* $v_1, \ldots, v_n$ *und* $w_1, \ldots, w_m$. *Dann gilt*

$$Rang\, A = \dim Im\, f.$$

Proposition 9.31 *Sind* $f : V \to V$ *und* $g : V \to V$ *zwei linearen Abbildungen mit Matrizen* A *und* B *in einer Basis, so ist* $A + B$ *(bzw.* AB *) ist die Matrix von* $f + g$ *(bzw. von* $f \circ g$*).*

9.3 Das charakteristische Polynom, Eigenwerte

Definition 9.32 Sei $A \in M(n \times n, K)$ eine quadratische Matrix. Dann heißt das Polynom

$$\chi_A(t) = |A - t \cdot E_n|$$

das charakteristische Polynom der Matrix A.

Beispiel 9.33 Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so ist das charakteristische Polynom von A zu

$$t^2 - (a + d)t + ad - bc$$

gleich.

Proposition 9.34 *Sei* $\varphi : V \to V$ *eine lineare Abbildung. Sind* $v_1, \ldots, v_n$ *und* $v'_1, \ldots, v'_n$ *zwei Basen von* V *und* A *(bzw.* A'*) die Matrix von* φ *in der Basis* $v_1, \ldots, v_n$ *(bzw. in* $v'_1, \ldots, v'_n$*), so ist*

$$\chi_A(t) = \chi_{A'}(t).$$

Dieses Polynom heißt das **charakteristisches Polynom** *von* φ *und wird mit* $\chi_\varphi(t)$ *bezeichnet.*

Beweis.

$$|A' - tE_n| = |T^{-1}AT - tE_n| = |T^{-1}(A - tE_n)T| = |T^{-1}||T||A - tE_n| = |A - tE_n|.$$

□

Definition 9.35 Sei $\varphi : V \to V$ eine lineare Abbildung. Man nennt $\lambda \in K$ **Eigenwert** von φ, wenn es einen von Null verschiedenen Vektor $v \in V$ gibt, so dass gilt

$$\varphi(v) = \lambda v, \quad v \neq 0.$$

In diesem Falle nennt man v heißt **Eigenvektor** (zum Eigenwert λ) von φ.

Satz 9.36 *$\lambda \in K$ ist ein Eigenwert von φ genau dann, wenn λ eine Nullstelle des charakteristischen Polynoms $\chi_\varphi(t)$ ist.*

Beweis. Ist $\varphi(v) = \lambda v$ und sind $x_1, \ldots, x_n$ Koordinaten von v in einer Basis $v_1, \ldots, v_n$, so ist

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Damit gilt

$$(A - \lambda E_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

Also besitzt das LGS mit der Matrix $A - \lambda E_n$ eine von Null verschiedene Lösung. Deswegen gilt $|A - \lambda E_n| = 0$.
Nun sei λ eine Nullstelle von $\chi_\varphi(t)$, d.h., $|A - \lambda E_n| = 0$. Dann besitzt das LGS mit der Matrix $A - \lambda E_n$ eine von Null verschiedene Lösung. Wir betrachten die Komponenten dieser Lösung als Koordinaten eines von Null verschiedenen Vektors $v \in V$. Damit erhalten wir $\varphi(v) = \lambda v$. Also ist λ ein Eigenwert.

□

9.4 Der Satz von Cayley-Hamilton

Satz 9.37 (Cayley-Hamilton) *Sei $\chi_A(t) = a_n + a_{n-1}t + \cdots a_0 t^n$ das charakteristische Polynom der Matrix A. Dann gilt*

$$\chi_A(A) := a_n E_n + a_{n-1}A + \cdots a_0 A^n = 0.$$

Beweis. Wegen 18.4 gilt

$$(A - tE_n) \cdot (A - tE_n)^{\#} = \chi_A(t) E_n$$

Wir bezeichnen $C(t) := (A - tE)^{\#}$ und setzen

$$\chi_A(t) = a_0 t^n + a_1 t^{n-1} + \cdots + a_{n-1} t + a_n.$$

Die Einträge $c_{ij}(t)$ von $C(t)$ sind Polynome in t vom Grad $\leq n-1$ (weil jeder Eintrag $c_{ij}(t)$ Determinante einer $(n-1) \times (n-1)$-Matrix ist). Wir schreiben $C(t)$ als

$$C(t) = C_0 t^{n-1} + C_1 t^{n-2} + \cdots + C_{n-1}$$

mit einigen Matrizen $C_0, C_1, \ldots, C_{n-1} \in M(n \times n, K)$. Aus der Gleichung

$$(A - tE_n)C(t) = \chi_A(t) E_n$$

folgt

$$-C_0 = a_0 E_n \quad (\text{Koeffizienten bei } \mathrm{t^n}),$$

$$-C_1 + AC_0 = a_1 E_n, \quad (\text{Koeffizienten bei } \mathrm{t^{n-1}})$$

$$-C_2 + AC_1 = a_2 E_n, \quad (\text{Koeffizienten bei } \mathrm{t^{n-2}})$$

$$\cdots$$

$$-C_{n-2} + AC_{n-3} = a_{n-2} E_n \quad (\text{Koeffizienten bei } \mathrm{t^2}),$$

$$-C_{n-1} + AC_{n-2} = a_{n-1} E_n \quad (\text{Koeffizienten bei t}),$$

$$AC_{n-1} = a_n E_n \quad (\text{Koeffizienten bei } \mathrm{t^0}).$$

Nun multiplizieren wir von links die erste Gleichung mit A^n, die zweite mit A^{n-1}, usw. die letzte mit $A^0 = E$. Durch Aufsummieren erhalten wir:

$$-A^n C_0 + A^{n-1}(-C_1 + AC_0) + A^{n-2}(-C_2 + AC_1) + \cdots$$

$$\cdots + A^2(-C_{n-2} + AC_{n-3}) + A(-C_{n-1} + AC_{n-2}) + AC_{n-1} = 0 =$$

$$= a_0 A^n + a_1 A^{n-1} + a_2 A^{n-2} + \cdots + a_{n-2} A^2 + a_{n-1} A + a_n E_n.$$

Damit ist $\chi_A(A) = a_0 A^n + a_1 A^{n-1} + a_2 A^{n-2} + \cdots + a_{n-2} A^2 + a_{n-1} A + a_n E_n$ die Nullmatrix.

□

9.5 Minimalpolynom

Definition 9.38 Sei $A \in M(n \times n; K)$. Ein normiertes Polynom $\mu_A(t) \in K[t]$ heißt **Minimalpolynom von** A, wenn $\mu_A(A) = 0$ und $\deg \mu_A(t) \leq \deg p(t)$ für jedes $p(t) \in K[t]$ mit $p(A) = 0$. Ist A die Matrix einer linearen Abbildung $\varphi : V \to V$ in einer Basis $v_1, \ldots, v_n$, so heißt $\mu_A(t)$ **Minimalpolynom** von φ.

Proposition 9.39 *Das Minimalpolynom einer Matrix $A \in M(n \times n; K)$ existiert und ist eindeutig bestimmt. Ist $p(t) \in K[t]$ ein Polynom mit $p(A) = 0$, so ist $p(t)$ durch das Minimalpolynom $\mu_A(t)$ teilbar. Insbesondere teilt $\mu_A(t)$ das charakteristische Polynom von A.*

Beweis. Die Existenz von μ_A folgt sofort aus dem Satz von Cayley-Hamilton.
Ist $p(t) \in K[t]$ ein Polynom mit $p(A) = 0$, so ist $\deg \mu_A(t) \leq \deg p(t)$ und durch Division mit Rest erhalten wir

$$p(t) = q(t)\mu_A(t) + r(t), \quad \deg r(t) < \deg \mu_A(t).$$

Andererseits, $0 = p(A) = r(A)$. Widerspruch.
Sind $\mu'_A(t)$ und $\mu_A(t)$ zwei Minimalpolynome, so ist $\mu'_A | \mu_A$ und $\mu_A | \mu'_A$. Da die beiden Polynome $\mu'_A(t)$ und $\mu_A(t)$ normiert sind, erhalten wir $\mu'_A(t) = \mu_A(t)$.

□

Proposition 9.40 *Das Minimalpolynom einer linearen Abbildung $\varphi : V \to V$ hängt nicht von der Basis $v_1, \ldots, v_n$ ab.*

Beweis. Sei A die Matrix von φ in einer Basis $v_1, \ldots, v_n$ und A' die Matrix von φ in einer anderen Basis $v'_1, \ldots, v'_n$. Wegen 9.29 gilt $A' = T^{-1}AT$. Wir bemerken, dass für jedes $k \geq 0$ gilt

$$(A')^k = \underbrace{T^{-1}ATT^{-1}AT \cdots T^{-1}AT}_{k} = T^{-1} \underbrace{A \cdot A \cdots A}_{k} T = T^{-1}A^kT.$$

Daraus folgt, dass

$$p(A') = T^{-1}p(A)T,$$

wobei $p(t) \in K[t]$ ein beliebiges Polynom ist. Deswegen ist $p(A')$ die Nullmatrix genau dann, wenn $p(A)$ die Nullmatrix. Insbesondere gilt

$$\mu_A(A) = \mu_A(A') = 0,$$

$$\mu_{A'}(A) = \mu_{A'}(A') = 0.$$

Aus $\mu_A(A') = 0$ folgt, dass $\mu_{A'}|\mu_A$ (s. 9.39). Aus $\mu_{A'}(A) = 0$ folgt, dass $\mu_A|\mu_{A'}$ (s. 9.39). Da die beiden Polynome $\mu'_A(t)$ und $\mu_A(t)$ normiert sind, erhalten wir $\mu'_A(t) = \mu_A(t)$.

□

9.6 Summen und direkte Summen

Definition 9.41 Seien $U_1, \dots U_k$ Unterräume eines Vektorraums V. Die **Summe** von $U_1, \dots U_k$ ist die Teilmenge $U_1 + \dots + U_k$ in V, die aus aller Summen $u_1 + \dots + u_k$ mit $u_i \in U_i$ besteht.

Satz 9.42 (Dimensionsformel für Summen) *Seien W_1, W_2 zwei endlichdimensionale Unterräume eines K-Vektorraums V. Dann gilt*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Beweis. Es sei $\{e_1, \dots, e_k\}$ eine Basis von $W_1 \cap W_2$. Wir ergänzen diese Basis bis auf Basen $\{e_1, \dots, e_k, f_1, \dots, f_m\}$ und $\{e_1, \dots, e_k, v_1, \dots, v_n\}$ von W_1 und W_2. Also es gilt $\dim W_1 = k + m$, $\dim W_2 = k + n$, $\dim W_1 \cap W_2 = k$. Nun zeigen wir, dass

$$e_1, \dots, e_k, f_1, \dots, f_m, v_1, \dots, v_n$$

eine Basis der Summe $W_1 + W_2$ ist.
Ist $x \in W_1 + W_2$ ein beliebiger Vektor, so gibt es $w_1 \in W_1$ und $w_2 \in W_2$ mit $x = w_1 + w_2$. Da w_1 (bzw. w_2) als Linearkombination von Vektoren $e_1, \dots, e_k, f_1, \dots, f_m$ (bzw. $e_1, \dots, e_k, v_1, \dots, v_n$) darstellbar ist, erhalten wir, dass x eine Linerkombination von

$$e_1, \dots, e_k, f_1, \dots, f_m, v_1, \dots, v_n$$

ist.
Es sei

$$0 = \alpha_1 e_1 + \dots + \alpha_k e_k + \beta_1 f_1 + \dots + \beta_m f_m + \gamma_1 v_1 + \dots + \gamma_n v_n.$$

Dann ist der Vektor

$$v := \alpha_1 e_1 + \dots + \alpha_k e_k + \beta_1 f_1 + \dots + \beta_m f_m$$

ein Element von W_1. Andererseits gilt

$$v = -\gamma_1 v_1 - \dots - \gamma_n v_n.$$

Damit ist v ein Vektor aus W_2. Also ist v ein Vektor aus $W_1 \cap W_2$. Da die Vektoren $e_1, \ldots, e_k$ eine Basis von $W_1 \cap W_2$ bilden, ist v als Linearkombination

$$v = \lambda_1 e_1 + \cdots + \lambda_k e_k$$

darstellbar. Damit erhalten wir

$$0 = v - v = \lambda_1 e_1 + \cdots + \lambda_k e_k + \gamma_1 v_1 + \cdots + \gamma_n v_n.$$

Aus der linearen Unabhängiggkeit von $e_1, \ldots, e_k, v_1, \ldots, v_n$ folgt, dass

$$\lambda_1 = \cdots = \lambda_k = \gamma_1 = \cdots = \gamma_n = 0.$$

Damit erhalten wir aus

$$0 = \alpha_1 e_1 + \cdots + \alpha_k e_k + \beta_1 f_1 + \cdots + \beta_m f_m + \gamma_1 v_1 + \cdots + \gamma_n v_n$$

die Gleichung

$$0 = \alpha_1 e_1 + \cdots + \alpha_k e_k + \beta_1 f_1 + \cdots + \beta_m f_m.$$

Nun folgt aus der linearen Unabhängiggkeit von $e_1, \ldots, e_k, f_1, \ldots, f_m$, dass

$$\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_m = 0.$$

Damit sind die Vektoren

$$e_1, \ldots, e_k, f_1, \ldots, f_m, v_1, \ldots, v_n$$

linear unabhängig und bilden eine Basis von $W_1 + W_2$. Deshalb gilt

$$\dim W_1 + W_2 = k + m + n = (k+m) + (k+n) - k = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

□

Definition 9.43 Seien $U_1, \ldots U_k$ Unterräume eines Vektorraums V. Die Unterräume $U_1, \ldots U_k$ heißen **unabhängig**, wenn aus $u_1 + \cdots + u_k = 0$ mit $u_i \in U_i$ $(i = 1, \ldots, k)$ stets $u_1 = u_2 = \cdots = u_k = 0$ folgt.

Proposition 9.44 *Seien $v_1, \ldots, v_k$ von Null verschiedene Vektoren in V. Die 1-dimensionalen Vektorräume $U_i := \mathrm{Span}(v_i)$ $(i = 1, \ldots, k)$ sind unabhängig genau dann, wenn die Vektoren $v_1, \ldots, v_k$ linear unabhängig sind.*

Beweis. Jeder Vektor $u_i \in U_i$ läßt sich als $u_i = \lambda_i v_i$ mit einem $\lambda_i \in K$ schreiben. Damit ist $u_1 + \cdots + u_k = 0$ zu $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$ äquivalent. Wegen $v_i \neq 0$ gilt $\lambda_i v_i = 0 \Leftrightarrow \lambda_i = 0$ $(i = 1, \ldots, k)$. □

Proposition 9.45 *Zwei Unterräume $U_1, U_2 \subset V$ sind unabhängig genau dann, wenn $U_1 \cap U_2 = 0$.*

Beweis. "$\Leftarrow$" : Ist $u_1 + u_2 = 0$ mit $u_1 \in U_1$ und $u_2 \in U_2$, so ist $u_1 = -u_2 \in U_1 \cap U_2 = 0$ und damit $u_1 = u_2 = 0$.
"$\Rightarrow$": Sind U_1 und U_2 unabhängig und $U_1 \cap U_2 \neq 0$ (d.h. es gibt $v \in U_1 \cap U_2$, $v \neq 0$), so kann man den Nullvektor als Summe $v + (-v)$ aufschreiben. Wegen $v \in U_1$ und $-v \in U_2$ erhalten wir einen Widerspruch. □

Definition 9.46 Seien $U_1, \dots U_k$ Unterräume eines Vektorraums V. Die Summe $U_1 + \cdots + U_k$ heißt **direkte Summe**, wenn $U_1, \dots U_k$ unabhängig sind. In diesem Fall schreibt man $U_1 \oplus \cdots \oplus U_k$.

Proposition 9.47 *Seien $U_1, \dots U_k$ Unterräume eines Vektorraums V und $W \subset V$ ein Unterraum. Dann gilt: $W = U_1 \oplus \cdots \oplus U_k$ genau dann, wenn jedes $w \in W$ eindeutig als Summe $w = u_1 + \cdots + u_k$ mit $u_i \in U_i$ darstellbar ist.*

Proposition 9.48 *Seien $U_1, \dots U_k$ Unterräume eines Vektorraums V und $v_1^{(i)}, \dots, v_{r_i}^{(i)}$ eine Basis von U_i $(i = 1, \dots, k)$. Dann gilt: die Summe $W := U_1 + \cdots + U_k$ ist direkt genau dann, wenn die Menge*

$$v_1^{(1)}, \dots, v_{r_1}^{(1)}, \dots, v_1^{(k)}, \dots, v_{r_k}^{(k)}$$

eine Basis von W ist.

Korollar 9.49 *Seien $U_1, \dots U_k$ Unterräume eines Vektorraums V. Sind $U_1, \dots U_k$ unabhängig, so gilt*

$$\dim(U_1 + \cdots + U_k) = \dim U_1 + \cdots + \dim U_k.$$

9.7 Diagonalisierbarkeit

Definition 9.50 Eine lineare Abbildung $\varphi : V \to V$ heißt **diagonalisierbar**, wenn es eine Basis $v_1, \dots, v_n$ von V gibt, so dass die Matrix $A = (a_{ij})$ von φ in der Basis $v_1, \dots, v_n$ Diagonalmatrix ist: d.h., $a_{ij} = 0 \ \forall i \neq j$.

Aus Definition der Matrix einer linearen Abbildung folgt sofort:

Proposition 9.51 *Sei A die Matrix von $\varphi : V \to V$ in einer Basis $v_1, \dots, v_n$. A ist eine Diagonalmatrix genau dann, wenn alle Vektoren $v_1, \dots, v_n$ Eigenvektoren sind.*

Proposition 9.52 *Sei $\varphi : V \to V$ eine lineare Abbildung. Sind $\lambda_1, \ldots, \lambda_k$ paarweise verschieden Elemente aus K, so sind die Unterräume $U_i := Ker\,\varphi - \lambda_i id$ $(i = 1, \ldots, k)$ unabhängig.*

Beweis. Induktion über k. Ist $k = 1$, so gibt es nichts zu beweisen. Nun sei $0 = u_1 + \cdots + u_k$ mit $u_i \in U_i$ $(i = 1, \ldots, k)$. Ist $u_i = 0$ für ein $i \in \{1, \ldots, k\}$, so kann man die Induktionsannnahme benützen und zeigen, dass alle Vektoren $u_1, \ldots, u_k$ gleich Null sind. Also werden wir behaupten, dass $u_i \neq 0 \ \forall \ i \in \{1, \ldots, k\}$. Andererseits gilt

$$0 = \varphi(0) = \varphi(u_1 + \cdots + u_k) = \lambda_1 u_1 + \cdots + \lambda_k u_k.$$

Damit erhalten wir:

$$0 = 0 - 0 = \lambda_k(u_1 + \cdots + u_k) - \lambda_1 u_1 + \cdots + \lambda_k u_k =$$

$$= (\lambda_1 - \lambda_k)u_1 + \cdots + (\lambda_1 - \lambda_{k-1})u_{k-1}.$$

Wegen der Induktionsannahme und $\lambda_i - \lambda_k \neq 0 \ \forall \ i \in \{1, \ldots, k-1\}$ erhalten wir $u_1 = \cdots = u_{k-1} = 0$. Damit ist auch $u_k = 0$.

□

Nun werden wir den Wert eines Polynoms $f(t) \in K[t]$ auf einem Endomorphismus $\varphi : V \to V$.

Definition 9.53 Sei V ein K-Vektorraum. Mit $\mathrm{End}(V)$ bezeichnen wir die Menge aller linearen Abbildungen $\varphi : V \to V$. Diese Menge besitzt die folgenden zwei Verknüpfungen

$$Summe: \quad (\varphi + \psi)(v) := \varphi(v) + \psi(v) \ \ \forall v \in V;$$

$$Produkt: \quad (\varphi \cdot \psi)(v) := \varphi(\psi((v)) \ \ \forall v \in V.$$

Es ist einfach zu zeigen, dass $\mathrm{End}(V)$ ein (nichtkommutativer) Ring mit Einselement ist. Wir nennen $\mathrm{End}(V)$ den **Endomorphismenring von** V. Ist $f(t) = a_0 + a_1 t + \cdots + a_k t^k \in K[t]$ ein Polynom, so definieren wir $f(\varphi)$ als

$$f(\varphi) := a_0 id + a_1 \varphi + \cdots + a_k \varphi^k.$$

Bemerkung 9.54 Ist A die Matrix von φ in einer Basis $v_1, \ldots, v_n$, so ist $f(A)$ die Matrix von $f(\varphi)$ in der Basis $v_1, \ldots, v_n$.

Bemerkung 9.55 Es ist wichtig zu bemerken, dass für je zwei Polynome $f(t), g(t) \in K[t]$ die Endomorphismen $f(\varphi)$ und $g(\varphi)$ miteinander kommutieren:

$$f(\varphi) \cdot g(\varphi) = g(\varphi) \cdot f(\varphi).$$

Diese Gleichung folgt aus der Gleichung

$$\varphi^i \cdot \varphi^j = \varphi^{i+j} = \varphi^j \cdot \varphi^i.$$

in der Berechnung

$$f(\varphi) \cdot g(\varphi) = \left(\sum_i a_i \varphi^i\right) \cdot \left(\sum_j b_j \varphi^j\right) = \sum_{ij} a_i b_j \varphi^{i+j} = \left(\sum_j b_j \varphi^j\right) \cdot \left(\sum_i a_i \varphi^i\right).$$

Proposition 9.56 *Sei $\varphi : V \to V$ eine lineare Abbildung. Es seien $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Elemente aus K, so dass gilt $f(\varphi) = 0$, wobei $f(t) := (t - \lambda_1) \cdots (t - \lambda_k)$ ist. Wir bezeichnen $U_i := Ker\, \varphi - \lambda_i id$ $(i = 1, \ldots, k)$. Dann gilt*

$$V = U_1 + \cdots + U_k.$$

Beweis. Für jedes $i \in \{1, \ldots, k\}$ definieren wir ein Polynom vom Grad $k - 1$:

$$g_i(t) := \frac{(t - \lambda_1) \cdots (t - \lambda_{i-1})(t - \lambda_{i+1}) \cdots (t - \lambda_k)}{(\lambda_i - \lambda_1) \cdots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \cdots (\lambda_i - \lambda_k)}.$$

Wir bemerken, dass $g_i(\lambda_i) = 1$ und $g_i(\lambda_j) = 0$ für $j \neq i$ ist. Aus der Interpolationsformel von Lagrange (s. 21.4) folgt:

$$g_1(t) + \cdots + g_k(t) = 1,$$

weil es nur ein einziges Polynom vom Grad $\leq k - 1$ gibt , welches den Wert 1 für $t \in \{\lambda_1, \ldots, \lambda_k\}$ besitz. Aber das Das Polynom $g_1(t) + \cdots + g_k(t)$ besitzt den Wert 1 für alle $t \in \{\lambda_1, \ldots, \lambda_k\}$ und $\deg(g_1(t) + \cdots + g_k(t)) \leq k - 1$. Daraus folgt, dass

$$id = g_1(\varphi) + \cdots + g_k(\varphi)$$

(man setzt φ in die letzte Gleichung ein). Ist $v \in V$ ein beliebiger Vektor, so erhalten wir

$$v = id(v) = g_1(\varphi)(v) + \cdots + g_k(\varphi)(v) = u_1 + \cdots + u_k,$$

wobei $u_i := g_i(\varphi)(v)$ $(i = 1, \ldots, k)$. Nun bemerken wir, dass $u_i \in U_i$ $\forall i \in \{1, \ldots, k\}$: wegen $f(\varphi) = (\varphi - \lambda_1 id) \cdots (\varphi - \lambda_k id) = 0$ erhalten wir

$$(\varphi - \lambda_i id)(u_i) = (\varphi - \lambda_i id) g_i(\varphi)(v) =$$

$$= \frac{(\varphi - \lambda_1 id) \cdots (\varphi - \lambda_k id)}{(\lambda_i - \lambda_1) \cdots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \cdots (\lambda_i - \lambda_k)}(v) = 0,$$

weil $F(\varphi) = \prod_{i=1}^{k} (\varphi - \lambda_i id)$ der Null-Endomorphismus ist. Damit haben wir gezeigt, dass $v \in U_1 + \cdots + U_k$.

□

Satz 9.57 (Kriterium der Diagonalisierbarkeit) *Sei V ein K-Vektorraum und $\varphi : V \to V$ eine lineare Abbildung. Die Abbildung φ ist diagonalisierbar genau dann, wenn das Minimalpolynom $\mu_\varphi(t)$ als Produkt $(t-\lambda_1)\cdots(t-\lambda_k)$ geschrieben werden kann, wobei $\lambda_1, \ldots, \lambda_k \in K$ und $\lambda_i \neq \lambda_j$ $\forall i \neq j$.*

Beweis. "⇒": Sei $\varphi : V \to V$ diagonalisierbar, d.h., es gibt eine Basis $v_1, \ldots, v_n$, so dass

$$\varphi(v_1) = a_{11}v_1, \ \ldots \ , \varphi(v_n) = a_{nn}v_n.$$

Die Einträge $a_{11}, a_{22}, \ldots, a_{nn}$ müssen nicht alle von einander verschieden sein. Sei $\{\lambda_1, \ldots, \lambda_k\}$ $(k \leq n)$ die Menge aller paarweise verschiedenen Werte von a_{ii} $i = 1, \ldots, n$. Wir zeigen, dass $f(t) := (t-\lambda_1)\cdots(t-\lambda_k)$ das Minimalpolynom von φ ist. Offensichtlich gilt

$$f(\varphi)(v_i) = 0 \ \ \forall i \in \{1, \ldots, n\},$$

weil jedes a_{ii} einem $\lambda_j \in \{\lambda_1, \ldots, \lambda_k\}$ gleich ist und

$$f(\varphi)(v_i) = f(a_{ii})v_i = 0 \cdot v_i = 0.$$

Damit gilt $f(\varphi) = 0$ und $f(t)$ ist durch das Minimalpolynom $\mu_\varphi(t)$ von φ teilbar. Andererseits gilt: Ist $\lambda \in K$ ein Eigenwert von φ, so ist λ eine Nullstelle von μ_φ, weil

$$\varphi(v) = \lambda v \ \Rightarrow \ 0 = \mu_\varphi(\varphi)v = \mu_\varphi(\lambda)v \ \Rightarrow \ \mu_\varphi(\lambda) = 0.$$

Deswegen ist $\mu_\varphi(t)$ durch jedes Polynom $(t-\lambda_i)$ $(i = 1, \ldots, k)$ teilbar. Damit ist $\mu_\varphi(t)$ durch $f(t)$ teilbar. Also gilt $\mu_\varphi(t) = f(t)$ (wegen $\mu_\varphi | f$ und $f | \mu_\varphi$).
"⇐": Sei $\mu_\varphi(t) := (t-\lambda_1)\cdots(t-\lambda_k)$ das Minimalpolynom von φ und $\lambda_i \neq \lambda_j$ $\forall i \neq j$. Wir definieren $U_i := Ker\,\varphi - \lambda_i id$ $(i = 1, \ldots, k)$. Aus 9.52 und 9.56 folgt, dass

$$V = U_1 \oplus \cdots \oplus U_k.$$

□

Definition 9.58 Eine lineare Abbildung $\varphi : V \to V$ heißt **Projektor**, wenn $\varphi^2 = \varphi$ gilt.

Korollar 9.59 *Jeder Projektor $\varphi : V \to V$ ist diagonalisierbar und es gilt*

$$V = Ker\,\varphi \oplus Im\,\varphi.$$

Beweis. Ein Projektor $\varphi : V \to V$ ist diagonalisierbar, weil das Minimalpolynom von φ das Polynom $t^2 - t = t(t-1)$ teilt. Es seien U_1 und U_0 Eigenräume zu den Eigenwerten 1 und 0. Dann gilt $V = U_1 \oplus U_0$. Offensichtlich gilt $U_0 = Ker\,\varphi$ und $U_1 \subset Im\,\varphi$. Wegen der Dimensionsformel erhalten wir $U_1 = Im\,\varphi$. □

Beispiel 9.60 Die lineare Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ mit der Matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

ist nicht diagonalisierbar. Wäre φ diagonalisierbar, so hätten wir in einer Basis eine Diagonalmatrix mit dem charakteristischen Polynom $(t-1)^2$. Daraus folgt, dass diese Matrix gleich der Einheitsmatrix wäre. Widerspruch (φ ist nicht identische Abbildung).

9.8 Invariante Unterräume

Definition 9.61 Sei $\varphi : V \to V$ eine lineare Abbildung und $U \subset V$ ein Unterraum. U heißt φ-invariant, wenn aus $v \in U$ stets $\varphi(v) \in U$ folgt.

Proposition 9.62 *Ist $\varphi : V \to V$ eine lineare Abbildung und $f(t) \in K[t]$ ein Polynom, so sind*

$$Ker\, f(\varphi) := \{v \in V \ : \ f(\varphi)(v) = 0\}$$

$$Im\, f(\varphi) := \{v \in V \ : \ \exists v' \in V \text{ mit } f(\varphi)(v') = v\}$$

φ-invariante Unterräume.

Beweis. Sei $v \in Ker\, f(\varphi)$. Dann gilt

$$f(\varphi)(\varphi(v)) = a_0\varphi(v) + a_1\varphi^2(v) + \cdots + \varphi^{k+1}(v) = \varphi \cdot (a_0 id + a_1\varphi + \cdots + a_k\varphi^k)(v) = 0$$

Damit erhalten wir $\varphi(v) \in Ker\, f(\varphi)$.
Sei $v \in Im\, f(\varphi)$ und $v = f(\varphi)(v')$. Dann gilt

$$\varphi(v) = \varphi(f(\varphi)(v')) = a_0\varphi(v') + a_1\varphi^2(v') + \cdots + \varphi^{k+1}(v') =$$

$$= (a_0 id + a_1\varphi + \cdots + a_k\varphi^k)(\varphi(v')) = f(\varphi)(\varphi(v')).$$

Also gilt $\varphi(v) \in Im\, f(\varphi)$

□

Korollar 9.63 *Ist $\varphi : V \to V$ eine lineare Abbildung, so sind*

$$Ker\, \varphi := \{v \in V \ : \ \varphi(v) = 0\}$$

$$Im\, \varphi := \{v \in V \ : \ \exists v' \in V \text{ mit } \varphi(v') = v\}$$

φ-invariante Unterräume.

Beweis. Die Aussage folgt aus 9.62 für $f(t) := t$. □

Proposition 9.64 *Sei $\varphi : V \to V$ eine lineare Abbildung und $U \subset V$ ein Unterraum. Ist $v_1, \ldots, v_n$ eine Basis von V, so dass $v_1, \ldots, v_k$ $(k \leq n)$ eine Basis von U, so gilt: der Unterraum U ist φ-invariant genau dann, wenn in der Matrix $A = (a_{ij})$ von φ in der Basis $v_1, \ldots, v_n$ alle Einträge a_{ij} mit $i > k$ und $j \leq k$ gleich Null sind:*

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

wobei $B \in Mat(k \times k, K)$, $D \in Mat(k \times (n-k), K)$, und $C \in Mat((n-k) \times (n-k), K)$.

Proposition 9.65 *Seien $U_1, \ldots U_k$ Unterräume eines Vektorraums V, $v_1^{(i)}, \ldots, v_{r_i}^{(i)}$ eine Basis von U_i $(i = 1, \ldots, k)$ und $\varphi : V \to V$ eine lineare Abbildung. Dann gilt: Sind alle Unterräume $U_1, \ldots U_k$ φ-invariant und $V := U_1 \oplus \cdots \oplus U_k$, so ist die Matrix A von φ in der Basis*

$$v_1^{(1)}, \ldots, v_{r_1}^{(1)}, \ldots, v_1^{(k)}, \ldots, v_{r_k}^{(k)}$$

eine Blockdiagonalmatrix:

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_k \end{pmatrix}$$

wobei jedes A_i eine $r_i \times r_i$-Matrix ist.

10 Polynome und ihre Ableitungen

Definition 10.1 Ist $f(t) = a_0 + a_1 t + \cdots + a_n t^n$ ein Polynom aus $K[t]$, so definiert man die Ableitung $f'(t)$ als

$$a_1 + 2a_2 t + \cdots + n a_n t^{n-1} \in K[t].$$

Proposition 10.2 *Für je zwei Polynome $f, g \in K[t]$ erhalten wir*
(i) $(f+g)' = f' + g'$;
(ii) $(f \cdot g)' = f'g + fg'$.

Proposition 10.3 *Es sei K ein Körper der Charakteristik 0. Ein Element $\lambda \in K$ ist eine Nullstelle von $f(t)$ der Vielfachheit m genau dann wenn gilt*

$$f(\lambda) = f'(\lambda) = \cdots = f^{(m-1)}(\lambda) = 0, \quad f^{(m)}(\lambda) \neq 0.$$

Beweis. $\lambda \in K$ ist eine Nullstelle von $f(t)$ der Vielfachheit m genau dann, wenn gilt $f(t) = (t-\lambda)^m g(t)$, wobei $g(\lambda) \neq 0$ ist. Wegen 10.2(ii) erhalten wir

$$f'(t) = m(t-\lambda)^{m-1} g(t) + (t-\lambda)^m g'(t) = (t-\lambda)^{m-1}(mg(t) + (t-\lambda)g'(t)).$$

Wir setzen $g_1(t) := mg(t) + (t-\lambda)g'(t)$. Dann gilt

$$f'(t) = (t-\lambda)^{m-1} g_1(t), \;\; g_1(\lambda) \neq 0$$

(wegen $g_1(\lambda) = mg(\lambda) \neq 0$). Also ist λ eine Nullstelle von $f(t)'$ der Vielfachheit $m-1$. Analog erhalten wir für $k \leq m$, dass λ eine Nullstelle von $f^{(k)}(t)$ der Vielfachheit $m-k$ ist. Daraus folgt die Aussage. □

Proposition 10.4 *Es seien $f, f_1, \ldots, f_k \in K[t]$ Polynome, so dass gilt*

$$ggT(f, f_i) = 1, \;\; i = 1, \ldots, k.$$

Dann gilt

$$ggT(f, f_1 \cdots f_k) = 1.$$

Beweis. Induktion über k. Der Fall $k = 1$ ist klar. Aus der Induktionsvoraussetzung folgt

$$ggT(f, f_1 \cdots f_{k-1}) = 1.$$

Es sei h ein gemeinsamer Teiler von f und $f_1 \cdots f_k$. Wegen $ggT(f, f_k) = 1$ gibt es Polynome $u, v \in K[t]$ mit

$$1 = uf + vf_k.$$

Durch Multiplikation mit $f_1 \cdots f_{k-1}$ erhalten wir

$$f_1 \cdots f_{k-1} = uf f_1 \cdots f_{k-1} + v f_1 \cdots f_{k-1} f_k.$$

Deswegen teilt h das Produkt $f_1 \cdots f_{k-1}$. Also ist h ein gemeinsamer Teiler von f und $f_1 \cdots f_{k-1}$. Wegen $ggT(f, f_1 \cdots f_{k-1}) = 1$ ist h eine Konstante. Damit erhalten wir

$$ggT(f, f_1 \cdots f_k) = 1.$$

□

Proposition 10.5 *Es seien $g, f_1, \ldots, f_k \in K[t]$ Polynome, so dass gilt*

$$ggT(f_i, f_j) = 1 \;\; \forall \;\; i \neq j$$

und f_i ein Teiler von g für $i = 1, \ldots, k$ ist. Dann ist das Produkt $f_1 \cdots f_k$ ein Teiler von g.

Beweis. Induktion über k. Der Fall $k = 1$ ist klar. Aus $ggT(f_k, f_i) = 1$ $(i = 1, \ldots, k-1)$ und 10.4 erhalten wir

$$ggT(f_k, f_1 \cdots f_{k-1}) = 1.$$

Deswegen existieren Polynome $u, v \in K[t]$ mit

$$1 = uf_k + vf_1 \cdots f_{k-1}.$$

Nach Induktionsvoraussetzung erhalten ist $f_1 \cdots f_{k-1}$ ein Teiler von g, d.h. $g = g_1 f_1 \cdots f_{k-1}$. Durch die Multiplikation mit g_1 erhalten wir

$$g_1 = uf_k g_1 + vg_1 f_1 \cdots f_{k-1} = uf_k g_1 + vg.$$

Also ist f_k ein Teiler von g_1, weil f_k das Polynom g teilt. Damit ist $f_1 \cdots f_k$ ein Teiler von $g = g_1 f_1 \cdots f_{k-1}$. □

Korollar 10.6 *Sind $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Elemente aus einem Körper K und $m_1, \ldots, m_r$ positive ganze Zahlen, so gilt:*
(i) *die Polynome $(t - \lambda_i)^{m_i}$ und $(t - \lambda_j)^{m_j}$ sind teilerfremd für $i \neq j$;*
(ii) *ist $(t - \lambda_i)^{m_i}$ ein Teiler eines Polynomes $g \in K[t]$ für $i = 1, \ldots, r$, so ist das Produkt*

$$(t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}$$

ein Teiler von g.

Satz 10.7 *Es seien $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Elemente aus einem Körper K und $m_1, \ldots, m_r$ positive ganze Zahlen mit $m_1 + \cdots + m_r = m$. Dann gibt es genau ein Polynom f vom Grad $\leq m - 1$, so dass gilt*

$$f^{(j)}(\lambda_i) = b_i^{(j)}, \;\; 1 \leq i \leq r, \;\; 0 \leq j \leq m_i - 1,$$

wobei

$$b_i^{(j)}, \;\; 1 \leq i \leq r, \;\; 0 \leq j \leq m_i - 1$$

gegebene Elemente des Körpers K sind. Insbesondere ist die $m \times m$-Matrix M des LGS für $a_0, \ldots, a_{m-1}$ ist nichtausgeartet (d.h. $\det M \neq 0$*).*

Beweis. Es sei

$$f(t) = a_0 + a_1 t + \cdots + a_{m-1} t^{m-1}$$

ein Polynom vom Grad $\leq n - 1$. Die $m = m_1 + \cdots + m_r$ Gleichungen

$$f^{(j)}(\lambda_i) = b_i^{(j)}, \;\; 1 \leq i \leq r, \;\; 0 \leq j \leq m_i - 1$$

definieren ein quadratisches LGS für die n unbekannten Koeffizienten $a_0, a_1, \ldots, a_{m-1}$. Dieses System besitzt eine einzige Lösung genau dann, wenn das assoziierte homogene Gleichungssystem nur die triviale Null-Lösung besitzt. Es sei $(a_0^{(0)}, a_1^{(0)}, \ldots, a_{m-1}^{(0)})$ eine Lösung des homogenen Gleichungssystems und

$$g(t) = a_0^{(0)} + a_1^{(0)} t + \cdots + a_{m-1}^{(0)} t^{m-1}$$

das entsprechende Polynom. Wegen

$$g^{(j)}(\lambda_i) = 0, \;\; 1 \leq i \leq r, \;\; 0 \leq j \leq m_i - 1$$

ist λ_i eine Nullstelle der Vielfachheit $\geq m_i$ von g $(i = 1, \ldots, r)$. Also ist $(t - \lambda_i)^{m_i}$ ein Teiler von g für $i = 1, \ldots, r$. Aus 10.6 folgt, dass das Produkt

$$(t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}$$

ein Teiler von g. Wegen $\deg g \leq m - 1$ und $m_1 + \cdots + m_r = m$ erhalten wir, dass g ein Nullpolynom sein muss. Damit ist $(a_0^{(0)}, a_1^{(0)}, \ldots, a_{n-1}^{(0)})$ eine triviale Lösung des homogenen Gleichungsystems. □

Beispiel 10.8 Ist $r = 1$ $m_1 = m$ $(\lambda_1 = \lambda, \; b_1^{(j)} = b^{(j)}, j = 0, 1, \ldots, m - 1)$, so bekommt man das Polynom f explizit mit Hilfe der Taylor-Formel

$$f(t) = b^{(0)} + b^{(1)}(t - \lambda) + \frac{b^{(2)}}{2!}(t - \lambda)^2 + \cdots + \frac{b^{(m-1)}}{(m-1)!}(t - \lambda)^{m-1}.$$

11 Funktionen von Matrizen

Satz 11.1 *Es sei $A \in M(n \times n, K)$ eine Matrix mit dem Minimalpolynom*

$$\mu_A(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r},$$

wobei $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Elemente von K und $m_1, \ldots, m_r \in \mathbb{Z}_{>0}$ $m = m_1 + \cdots + m_r$. Ist $f(t) \in K[t]$ ein Polynom, so gilt

$$f(A) = r(A),$$

wobei $r(t) \in K[t]$ das einzige Polynom vom Grad $\leq m - 1$ mit der Eigenschaft

$$r^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \;\; 1 \leq i \leq r, \;\; 0 \leq j \leq m_i - 1$$

ist.

Beweis. Es sei $r(t)$ der Rest von f nach Division mit $\mu_A(t)$. Dann gilt

$$f(t) = q(t)\mu_A(t) + r(t)$$

und

$$f(A) = r(A).$$

Offensichtlich gilt $f(\lambda_i) = r(\lambda_i)$ für $i = 1, \ldots, r$. Nach Differenzierung erhalten wir auch

$$f^{(j)}(\lambda_i) = r^{(j)}(\lambda_i), \;\; 1 \leq i \leq r, \;\; 0 \leq j \leq m_i - 1.$$

Damit bekommt man ein LGS für die Koeffizienten des Polynoms

$$r(t) = r_0 + r_1 t + \cdots + r_{m-1} t^{m-1}.$$

□

Beispiel 11.2 Es sei $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$. Dann ist $\mu_A(t) = (t-2)^2$. Zur Berechnung der Matrix A^n muss man das Restpolynom $r(t) = r_0 + r_1 t$ finden, wobei

$$t^n = q(t)\mu_A(t) + r(t).$$

Für die Koeffizienten r_0, r_1 erhalten wir das LGS:

$$2^n = r(2) = r_0 + r_1 2,$$

$$n2^{n-1} = r'(2) = r_1.$$

Also gilt $r_1 = n2^{n-1}$ und $r_0 = 2^n - n2^n = (1-n)2^n$. Damit ergibt sich

$$A^n = r(A) = r_0 E_2 + r_1 A = (1-n)2^n \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + n2^{n-1} \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} =$$

$$= 2^{n-1} \begin{pmatrix} 2-n & n \\ -n & 2+n \end{pmatrix}.$$

Definition 11.3 Es sei

$$f(t) = \sum_{l \geq 0} c_l t^l, \;\; c_l \in K$$

eine Potenzreihe und $A \in M(n \times n, K)$ $(K = \mathbb{R}, \mathbb{C})$ eine Matrix. Wir bezeichnen

$$f_k(t) := \sum_{l=0}^{l} c_l t^l$$

und definieren $f(A)$ als

$$f(A) := \lim_{k \to \infty} f_k(A)$$

(falls $\lim_{k\to\infty} f_k(A)$ existiert).

Satz 11.4 *Es sei $A \in M(n \times n, K)$ $(K = \mathbb{R}, \mathbb{C})$ eine Matrix mit dem Minimalpolynom*

$$\mu_A(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}.$$

Ist

$$f(t) = \sum_{l \geq 0} c_l t^l, \quad c_l \in K$$

eine Potenzreihe, so dass die Ableitungen

$$f^{(j)}(\lambda_i) = \sum_{l \geq 0} c_l l(l-1) \cdots (l-j-1) \lambda^{l-j}$$

absolut konvergieren für $1 \leq i \leq r, \quad 0 \leq j \leq m_i - 1$. Dann existiert $f(A)$. Die Matrix $f(A)$ läßt sich als $r(A)$ berechnen, wobei $r(t)$ ein Polynom vom Grad $\leq m-1$ $(m = m_1 + \cdots + m_r)$ mit der Eigenschaft

$$r^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \quad 1 \leq i \leq r, \quad 0 \leq j \leq m_i - 1$$

ist.

Beweis. Wir dividieren $f_k(t) = \sum_{l \geq 0}^{k} c_l t^l$ durch $\mu_A(t)$ mit Rest:

$$f_k(t) = q_k(t)\mu_A(t) + r_k(t).$$

Für die Koeffizienten des Polynoms

$$r_k(t) = r_0^{(k)} + r_1^{(k)} t + \cdots + r_{m-1}^{(k)} t^{m-1}$$

erhalten wir LGS aus der Bedingungen

$$r_k^{(j)}(\lambda_i) = f_k^{(j)}(\lambda_i), \quad 1 \leq i \leq r, \quad 0 \leq j \leq m_i - 1.$$

Es sei M die Matrix des LGS. Dann ist M unabhängig von k und $|M| \neq 0$ (s. 10.7). Aus

$$M \begin{pmatrix} r_0^{(k)} \\ r_1^{(k)} \\ \cdot \\ r_{n-1}^{(k)} \end{pmatrix} = \begin{pmatrix} \cdot \\ f_k^{(j)}(\lambda_i) \\ \cdot \\ \cdot \end{pmatrix}$$

folgt

$$\begin{pmatrix} r_0^{(k)} \\ r_1^{(k)} \\ \cdot \\ r_{n-1}^{(k)} \end{pmatrix} = M^{-1} \begin{pmatrix} \cdot \\ f_k^{(j)}(\lambda_i) \\ \cdot \\ \cdot \end{pmatrix}.$$

Deswegen existieren die Grenzwerte

$$r_i := \lim_{k\to\infty} r_i^{(k)}, \quad i = 0, 1, \ldots, n-1,$$

weil

$$\lim_{k\to\infty} f_k^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \quad 1 \le i \le r, \quad 0 \le j \le m_i - 1$$

existiert. Also existiert auch

$$\lim_{k\to\infty} f_k(A) = \lim_{k\to\infty} r_k(A) = (\lim_{k\to\infty} r_0^{(k)})E_n + (\lim_{k\to\infty} r_1^{(k)})A + \cdots + (\lim_{k\to\infty} r_{m-1}^{(k)})A^{m-1} =$$

$$= r_0 E_n + r_1 A + \cdots + r_{m-1}A^{m-1} = r(A),$$

wobei die Koeffizienten von $r(t)$ Lösung des LGS

$$M \begin{pmatrix} r_0 \\ r_1 \\ \cdot \\ r_{m-1} \end{pmatrix} = \begin{pmatrix} \cdot \\ f^{(j)}(\lambda_i) \\ \cdot \\ \cdot \end{pmatrix}$$

sind. □

Beispiel 11.5 Es sei $A = \begin{pmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{pmatrix}$. Dann ist $\mu_A(t) = (t-2)(t-1)$. Zur Berechnung der Matrix e^A muss man das Polynom $r(t) = r_0 + r_1 t$ finden. Die Koeffizienten r_0, r_1 sind definiert durch das LGS

$$e^{\lambda_1} = r(\lambda_1) = r_0 + r_1\lambda_1,$$

$$e^{\lambda_2} = r(\lambda_2) = r_0 + r_1\lambda_2,$$

wobei $\lambda_1 = 1$,$\lambda_2 = 2$. Damit erhalten wir:

$$e = r_0 + r_1,$$

$$e^2 = r_0 + r_1 2.$$

Also gilt $r_1 = e^2 - e$ und $r_0 = e - (e^2 - e) = 2e - e^2$. Es ergibt sich

$$e^A = r(A) = r_0 E_2 + r_1 A = (2e - e^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (e^2 - e)\begin{pmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{e^2+e}{2} & \frac{e^2-e}{2} \\ \frac{e^2-e}{2} & \frac{e^2+e}{2} \end{pmatrix}.$$

12 Rekurrente Folgen

Definition 12.1 Sei K ein Körper. Eine eine unendliche Folge $x = \{x_0, x_1, x_2, \ldots, x_i, \ldots\}$ von Elementen aus K heißt **rekurrente Folge**, wenn eine positive Zahl d und Elemente $a_1, a_2, \ldots, a_d \in K$ gibt, so dass die folgende **rekurrente Gleichung** gilt

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_d x_{n-d}, \quad \forall n \geq d.$$

Beispiel 12.2 Die Folge der Fibonaccischen Zahlen:

$$\{1, 1, 2, 3, 5, 8, 13, 21, \ldots\}$$

$x_0 = 1, x_1 = 1, x_n = x_{n-1} + x_{n-2}$ $(n \geq 2)$ ist eine rekurrente Folge, wobei $a_1 = a_2 = 1$

Satz 12.3 *Die Menge $R(a_1, \ldots, a_d)$ aller rekurrente Folgen $\{x_i\}_{i\geq 0}$ mit einer gegebenen rekurrenten Relation*

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_d x_{n-d}, \quad \forall n \geq d$$

ist ein K-Vektorraum der Dimension d, wobei Addition und Multiplikation mit Skalar wie folgt definiert ist:

$$\{x_i\}_{i\geq 0} + \{y_i\}_{i\geq 0} := \{x_i + y_i\}_{i\geq 0},$$

$$\lambda\{x_i\}_{i\geq 0} := \{\lambda x_i\}_{i\geq 0}.$$

Beweis. Zunächst zeigen wir:
(1) Ist $x = \{x_i\}_{i\geq 0}, y = \{y_i\}_{i\geq 0} \in R(a_1, \ldots, a_d)$, so ist $x + y = \{x_i + y_i\}_{i\geq 0} \in R(a_1, \ldots, a_d)$.
(2) Ist $x = \{x_i\}_{i\geq 0} \in R(a_1, \ldots, a_d)$, so ist $\lambda x = \{\lambda x_i\}_{i\geq 0} \in R(a_1, \ldots, a_d)$.
Daraus folgt, dass $R(a_1, \ldots, a_d)$ ein K-Vektorraum ist.
Nun bestimmen wir eine Basis $e_0, \ldots, e_{d-1}$ von $R(a_1, \ldots, a_d)$: Sei e_k $(0 \leq k \leq d-1)$ eine Folge $\{z_i\}_{i\geq 0} \in R(a_1, \ldots, a_d)$, die durch die folgenden Anfangsbedingungen eindeutig definiert ist:

$$z_i = \begin{cases} 0, & \text{wenn } i \neq k \text{ und } 0 \leq i \leq d-1 \\ 1, & \text{wenn } i = k \end{cases}$$

$$e_0 = \{\underbrace{1, 0, \ldots, 0}_{d}, *, *, \ldots\},$$

$$e_1 = \{\underbrace{0, 1, \ldots, 0}_{d}, *, *, \ldots\},$$

$$\dots$$

$$e_{d-1} = \{\underbrace{0, 0, \dots, 1}_{d}, *, *, \dots\}.$$

Die Elemente $e_0, \dots, e_{d-1} \in R(a_1, \dots, a_d)$ sind linear unabhängig, weil die ersten d Komponenten von $e_0, \dots, e_{d-1}$ Zeilen der Einheitsmatrix E_d sind. Andererseits ist jede Folge $x = \{x_i\}_{i\geq 0} \in R(a_1, \dots, a_d)$ gleich der Linearkombination $x_0e_1 + \dots + x_{d-1}e_{d-1}$, weil in der Folge $y = \{y_i\}_{i\geq 0} := x - x_0e_1 + \dots + x_{d-1}e_{d-1} \in R(a_1, \dots, a_d)$ die ersten d Komponenten $y_0, y_1, \dots, y_{d-1}$ gleich Null sind, und wegen der rekurrenten Relation

$$y_n = a_1y_{n-1} + a_2y_{n-2} + \dots + a_dy_{n-d}, \quad \forall n \geq d$$

alle Komponenten von y gleich Null sind. Also ist $e_0, \dots, e_{d-1}$ eine Basis von $R(a_1, \dots, a_d)$ und damit $\dim_K R(a_1, \dots, a_d) = d$. □

Proposition 12.4 *Sei λ ein Element aus K. Die Folge*

$$e(\lambda) := (1, \lambda, \lambda^2, \dots, \lambda^k, \dots)$$

ist ein Element von $R(a_1, \dots, a_d)$ genau dann, wenn λ eine Nullstelle des Polynoms

$$f(t) = t^d - a_1t^{d-1} - a_2t^{d-2} - \dots - a_d$$

ist. Das Polynom $f(t)$ wird **charakteristisches Polynom** *der Rekurrenzrelation in $R(a_1, \dots, a_d)$ genannt.*

Beweis. Ist $e(\lambda) \in R(a_1, \dots, a_d)$, so ist

$$\lambda^d = a_1\lambda^{d-1} + a_2\lambda^{d-2} + \dots + a_d \cdot 1.$$

Damit ist λ eine Nullstelle von $f(t)$.
Ist λ eine Nullstelle von $f(t)$, so erhalten wir durch Multiplikation mit λ^{n-d}:

$$\lambda^n = a_1\lambda^{n-1} + a_2\lambda^{n-2} + \dots + a_d\lambda^{n-d}.$$

Deswegen gilt $e(\lambda) \in R(a_1, \dots, a_d)$. □

Proposition 12.5 *Sei λ eine Nullstelle der Vielfachheit m des Polynoms*

$$f(t) = t^d - a_1t^{d-1} - a_2t^{d-2} - \dots - a_d.$$

Dann sind die Folgen

$$e(\lambda) := (1, \lambda, \lambda^2, \lambda^3, \dots, \lambda^k, \dots)$$

$$e'(\lambda) := (0, 1, 2\lambda, 3\lambda^2, \ldots, k\lambda^{k-1}, \ldots)$$
$$e''(\lambda) := (0, 0, 2, 6\lambda, 12\lambda^2, \ldots, k(k-1)\lambda^{k-2}, \ldots)$$
$$\cdots$$
$$e^{(m-1)}(\lambda) := (0, 0, \cdots, \underbrace{(m-1)!}_{m}, m(m-1)\cdots 2\lambda, \ldots, k(k-1)\cdots(k-m+2)\lambda^{k-m+1}, \ldots)$$

Elemente von $R(a_1, \ldots, a_d)$.

Beweis. Da das Polynom $t^{n-d}f(t)$ die Nullstelle λ der Vielfachheit $\geq m$ besitzt, erhalten wir für $j = 0, 1, \ldots, m-1$ und für alle $n \geq d$:

$$(t^{n-d}f(t))^{(j)}(\lambda) = 0 = n(n-1)\cdots(n-j+1)\lambda^{n-j} - a_1(n-1)(n-2)\cdots(n-j)t^{n-j-1} -$$
$$-a_2(n-2)(n-3)\cdots(n-j-1)t^{n-j-2} - \cdots - a_d(n-d)(n-d-1)\cdots(n-d-j+1)\lambda^{n-d-j},$$

wobei wir $\lambda^k = 0$ für $k < 0$ setzen. Daraus folgt, dass

$$e^{(j)}(\lambda) := (0, 0, \cdots, \underbrace{j!}_{j+1}, (j+1)j\cdots 2\lambda, \ldots, k(k-1)\cdots(k-j+1)\lambda^{k-j}, \ldots,$$

ein Element von $R(a_1, \ldots, a_d)$ für $j = 0, 1, m-1$. □

Satz 12.6 *Sei K ein Körper. Besitzt das Polynom*

$$f(t) = t^d - a_1 t^{d-1} - a_2 t^{d-2} - \cdots - a_d$$

Nullstellen $\lambda_1, \ldots, \lambda_l$ der Vielfachheit $m_1, \ldots, m_l$, wobei $m_1 + \cdots + m_l = d$, so ist

$$e(\lambda_1), e'(\lambda_1), \ldots, e^{(m_1-1)}(\lambda_1), \ldots, e(\lambda_l), e'(\lambda_l), \ldots, e^{(m_l-1)}(\lambda_l)$$

eine Basis von $R(a_1, \ldots, a_d)$.

Beweis. Wegen $\dim R(a_1, \ldots, a_d) = d$ reicht es zu zeigen, dass die Folgen

$$e(\lambda_1), e'(\lambda_1), \ldots, e^{(m_1-1)}(\lambda_1), \ldots, e(\lambda_l), e'(\lambda_l), \ldots, e^{(m_l-1)}(\lambda_l)$$

linear unabhängig sind. Dies folgt aus der Tatsache, dass die ersten d Elemente dieser Folgen eine nichtausgeartete $d \times d$-Matrix bilden (s. 10.7). □

Korollar 12.7 *Besitzt das Polynom*

$$f(t) = t^d - a_1 t^{d-1} - a_2 t^{d-2} - \cdots - a_d$$

genau d verschiedenen Nullstellen $\lambda_1, \ldots, \lambda_d$, so ist

$$e(\lambda_1), e(\lambda_2), \ldots, e(\lambda_d)$$

eine Basis von $R(a_1, \ldots, a_d)$.

Beispiel 12.8 Das charakteristische Polynom der rekurrenten Folgen in $R(1,1)$ ist $t^2 - t - 1$. Dieses Polynom hat zwei verschiedene Nullstellen

$$\lambda_1 = \frac{1+\sqrt{5}}{2},\ \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

Damit ist Fibonaccische Folge als Linearkombination $a \cdot e(\lambda_1) + b \cdot e(\lambda_2)$ darstellbar. Um die Koeffizienten a und b zu bestimmen, haben wir die Gleichungen

$$f_0 = 1 = a + b,$$

$$f_1 = 1 = a\frac{1+\sqrt{5}}{2} + b\frac{1-\sqrt{5}}{2}.$$

Daraus folgt, dass $a - b = 1/\sqrt{5}$ und

$$a = \frac{1+1/\sqrt{5}}{2},\ \ b = \frac{1-1/\sqrt{5}}{2}.$$

Damit bekommen wir eine explizite Formel für Fibonaccische Zahlen

$$f_k = \frac{1+1/\sqrt{5}}{2}\left(\frac{1+\sqrt{5}}{2}\right)^k + \frac{1-1/\sqrt{5}}{2}\left(\frac{1-\sqrt{5}}{2}\right)^k =$$

$$= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{k+1}.$$